

The transcriptome of the marine calanoid copepod *Temora longicornis* under heat stress and recovery

Ilias Semmouri^a, Jana Asselman^a, Filip Van Nieuwerburgh^b, Dieter Deforce^b, Colin R. Janssen^a, Karel A.C. De Schamphelaere^a

^a Ghent University, Laboratory of Environmental Toxicology and Aquatic Ecology, Faculty of Bioscience Engineering, 9000 Ghent, Belgium

^b Ghent University, Laboratory of Pharmaceutical Biotechnology, Faculty of Pharmaceutical Sciences, 9000 Ghent, Belgium

contact corresponding author: ilias.semmouri@ugent.be

Present address: Ghent University, Laboratory of Environmental Toxicology and Aquatic Ecology, Campus Coupure, Building F – 2nd Floor, Coupure Links 653 – B-9000 Ghent – Belgium

Abstract

Understanding the impacts of global change in zooplankton communities is crucial, as alterations in the zooplankton communities can affect entire marine ecosystems. Despite the economic and ecological importance of the calanoid copepod *Temora longicornis* in the Belgian part of the North Sea, molecular data is still very limited for this species. Using HiSeq Illumina sequencing, we sequenced the whole transcriptome of *T. longicornis*, after being exposed to realistic temperatures of 14 and 17 °C. After both an acute (1 day) and a more sustained (5 days) thermal exposure to 17 °C, we investigated gene expression differences with animals exposed to 14 °C, which may be critical for the thermal acclimation and resilience of this copepod species. We also studied the possibility of a short term stress recovery of a heat shock. A total of 179,569 transcripts were yielded, of which 44,985 putative ORF transcripts were identified. These transcripts were subsequently annotated into roughly 22,000 genes based on known sequences using Gene Ontology (GO) and KEGG databases. *Temora* only showed a mild response to both the temperature and the duration of the exposure. We found that the expression of 27 transcripts varied significantly with an increase in temperature of 3 °C, of which eight transcripts were differentially expressed after acute exposure only. Gene set enrichment analysis revealed that, overall, *T. longicornis* was more impacted by a sustained thermal exposure, rather than an immediate (acute) exposure, with two times as many enriched GO terms in the sustained treatment. We also identified several general stress responses independent of exposure time, such as modified protein synthesis, energy mobilisation, cuticle and chaperone proteins. Finally, we highlighted candidate genes of a possible recovery from heat exposure, identifying similar terms as those enriched in the heat treatments, i.e. related to for example energy metabolism, cuticle genes and extracellular matrix. The data presented in this study provides the first transcriptome available for *T. longicornis* which can be used for future genomic studies.

Keywords: North Sea, BPNS, global warming, global change, climate change, temperature, gene expression, Crustacea, Arthropoda, zooplankton

1. Introduction

Over the past decades, the world's oceans and seas have been influenced by human induced impacts such as overfishing, invasive species, coastal eutrophication, chemical pollution, oil/gas extraction and global change (Crain et al., 2009; Halpern et al., 2008; Zacharias and

60 Gregr, 2005). Globally about 41 % of our oceans and seas are already impacted to a high or
61 very high degree, including several parts of the North Sea (Halpern et al., 2008). A major aspect
62 of global change is global warming, i.e. the long-term increase in the Earth's average surface
63 temperature. In the North Sea region, which is characterized by large annual fluctuations in the
64 temperature of the surface water (from ca. 5°C in winter to about 20 °C in summer and an
65 annual mean SST between 10.5 °C and 11.5°C), the average sea surface temperature (SST) of
66 the water has already risen with 1-2 °C over a time-period of twenty-five years (1985 - 2010)
67 (OSPAR, 2010; van der Kooij et al., 2016). In comparison with the warming rate of the western
68 part of the North-east Atlantic (0.4 - 0.8 °C) over the same time-period, the North Sea is
69 warming faster (OSPAR, 2010). Moreover, the global SST is likely to rise further within the
70 range of 1.1–6.4 °C during the 21st century, depending on the selected emission scenario (Aral
71 and Guan, 2016; IPCC, 2007).

72 Over 99% of all organisms on Earth are ectothermic, making them susceptible to fluctuations
73 of temperature, and hence, to global warming (Atkinson, 1994). Furthermore, an increased
74 temperature also affects the physical and chemical properties of water, e.g. solubility of carbon
75 dioxide (CO₂) and oxygen (O₂) and the salinity, viscosity, pH and the density of the sea water
76 (IPCC, 2014; McNeil and Matear, 2006; Orr et al., 2005; Pierce et al., 2012; Sharqawy et al.,
77 2010). From a food web point of view, primary production by the smallest plankton is expected
78 to increase in the warmer waters, but this often does not translate into secondary production
79 (the zooplankton and smaller fish) which shows a decreased productivity due to the with global
80 change associated ocean acidification (Nagelkerken and Connell, 2015). As a consequence,
81 temperature often has striking impacts on virtually all life processes from a molecular scale to
82 the entire pelagic ecosystem scale and from plankton to higher trophic levels (Beaugrand and
83 Kirby, 2010; Beaugrand et. al., 2009).

84 Temperature is considered the major factor structuring marine ecosystems and shaping
85 zooplankton dynamics in particular (Richardson 2008). Due to their rapid responses to
86 environmental variation, including shifts in temperature (i.e. changes in biomass, community
87 structure and trophic linkages in marine food webs), planktonic organisms are frequently used
88 as bio-indicators of ecosystem changes (Chiba et al., 2018; Ferdous and Muktadir, 2009).
89 Therefore, understanding the impacts of this changing environmental condition, and hence
90 global change, is essential, as alterations in the zooplankton communities can affect entire
91 marine food webs and ecosystems (Richardson 2008). Zooplanktonic abundance in the Belgian
92 Part of the North Sea (BPNS) is dominated year-round by copepods (on average around 66 %

and sometimes even up to 84 %) (Krause et al. 1995; Van Ginderdeuren, 2014b). Copepods are the most important trophic link between primary producers and higher-level consumers (such as fish and marine mammals) in the pelagic food web (Möllmann et al., 2005; Richardson, 2008; Van Ginderdeuren, 2014a). By representing 35 to 70% of the total copepod community, the calanoid copepod *Temora longicornis* O.F. Müller (Maxillopoda, Copepoda, Calanoida) is a major component of the copepod assemblage in the coastal waters of the northern hemisphere (Brylinski, 2009; Daan, 1989; Halsband and Hirche, 2001; Van Ginderdeuren, 2014b; Williams et al., 1994). *T. longicornis* has been shown to be able to remove up to 49% of the daily primary production of phytoplankton, illustrating its ecological importance (Dam and Peterson, 1993). This copepod species is an important primary food source for the pelagic fish stocks of the BPNS, such as the commercially important herring, *Clupea harengus* L. (Casini et. al., 2004; Van Ginderdeuren et. al., 2014a).

Due to its ecological and economic importance, the impact of temperature (and climate change) on this species has already been investigated numerous times. Several studies suggest that abundance of *T. longicornis* is strongly correlated to climate variability (Deschutter et al., 2017; Dippner et al., 2000; Heyen et al., 1998; Kane and Prezioso, 2008). It was found that *T. longicornis* was not able to withstand temperatures higher than 22.5 °C, explaining its absence in warmer environments and its restriction to the northern hemisphere, with a southern boundary coinciding with the 20 °C isotherm in summer (Halsband-Lenk et. al., 2002). Many studies have already reported effects of increased temperature on different life history traits of this species including body size, development and a number of reproductive parameters (Devreker et al., 2005; Dzierzbicka-Glowacka et al., 2011; Halsband-Lenk et al., 2002; Hay et al., 1988; Holste et. al., 2009; Klein Breteler and Gonzalez, 1986). Together, the results of all of the studies mentioned above, indicate that *Temora longicornis* can be regarded as a sentinel species for assessing the impact of global change on the marine environment. Despite the abundance, global distribution and importance of copepods, only limited biological information is available at the molecular level for this taxon (e.g. Lee et. al., 2015; Lenz et. al., 2014; Ning et. al., 2013). Nevertheless, molecular studies on the physiological responses to stress are an interesting method to study and predict responses to climate change, since responses of life-history parameters to environmental stress are being coordinated by changes in gene networks and pathways (Bron et. al., 2011). For *T. longicornis* in particular, no genomic/transcriptomic studies have been reported so far. To date, there are only 31 nucleotide sequences for *T. longicornis* deposited in GenBank (September, 2018), corresponding to COI and ribosomal

genes. Global gene expression could identify cellular and molecular processes, which may provide key data on the effect that global warming has and will have on this abundant species at higher levels of organisation at the individual, population and community level. Therefore, in this study, we used RNA sequencing technology to study the transcriptomics responses and mechanisms of the calanoid copepod *T. longicornis* to temperature fluctuations. Our objective is to document gene expression changes in response to different thermal stress regimes in this ecologically relevant marine species. Using realistic temperatures (i.e. 14 °C, the spring and summer mean SST of the BPNS and 17°C, an estimation of the average SST of the next century), we simulated an increase of temperature of 3 °C for different durations to find genetic mechanisms and signatures for both an immediate or short term (acute) and a more sustained (subchronic) exposure. Another aspect of stress responses in invertebrates that has received relatively little attention is stress recovery. Understanding the mechanisms and processes enabling recovery from stress will help to shed light on the occurring processes after stress exposure. We, therefore, investigated the possibility of a short-term recovery of a heat shock, aiming to identify signatures of stress recovery.

2. Material & Methods

2.1. Sampling and animal acclimatization

Zooplankton were collected with the research vessel (RV) Simon Stevin on September 24th and 25th 2015, from the coastal waters near Ostend (Belgium), North Sea, NO Atlantic Ocean (September 24: Lat: 51° 19' 9.62" N, Long: 2° 30' 52" E, September 25: Lat: 51° 28' 5.55" N, Long: 2° 48' 27.66" E). Copepods were collected using a WP2 net (200 µm mesh size) and specimens were stored in 30 L vessels until they could be transported to the laboratory. There, live *T. longicornis* individuals were picked out with a Pasteur pipette and transferred into a rigorously aerated 35 L vessel containing filtered natural seawater (FSW) from the sampling location (Lat: 51° 28' 5.55" N, Long: 2° 48' 27.66" E, sea surface), at a temperature of 14 °C, a salinity of 34.05 PSU and a conductivity of 4.3 S/m.

During the acclimation period, copepods were fed three times a week *ad libitum* with a 3:1 mixture of live algae (cf. Gonçalves et al., 2014; Klein Breteler, 1980), based on cell densities, containing respectively 75% *Rhodomonas salina* (Wislouch) D.R.A.Hill & R.Wetherbee (Cryptophyta) and 25% of the other algae species: *Isochrysis galbana* Parke (Haptophyta), *Oxhyrris marina* Dujardin (Dinoflagellata), *Prorocentrum cordatum* (Ostenfeld) J.D.Dodge (Dinoflagellata) and *Thalassiosira weissflogii* (Grunow) G.Fryxell & Hasle (Bacillariophyta).

After seven days of acclimation, adult (fully developed) *T. longicornis* were selected from the culture for the exposure experiment.

2.2. Experimental design

Adult (fully developed) *T. longicornis* individuals were first exposed to a temperature of either 14 °C or 17 °C. A pool of both male and female copepods was used in the experiment in a sex ratio of 7:10 males:females, which was the same sex ratio that we observed in the zooplankton sample. Temperatures were chosen based on 1) the long-term spring/summer mean SST (April to September) of the southern part of the North Sea i.e. 14 °C (Klimaatinfo, 2016) and 2) a temperature three degrees higher, representing an estimation of the average temperature of the next century if global warming is not stopped or slowed down in the next hundred years (IPCC, 2007). Moreover, as mentioned above, the temperature of the surface water of the BPNS can reach to 20°C in summer, indicating the chosen temperatures are already relevant today. Eight replicates were kept for four days at 14 °C and the other eight were acclimatized to 17 °C over the course of 24 h and kept at this temperature for the remaining three days (Figure 1). Subsequently at day four, four randomly selected replicates were moved for 24 hours to 17 °C from the 14 °C group and vice versa. This resulted in four treatment groups (1) a group exposed to 14°C for five days (2) a group exposed to 17°C for five days (below referred to as the ‘sustained exposure’ group) (3) a group exposed to 14°C for four days and then to an additional acute exposure of 17°C (below referred to as acute exposure group) and (4) a group exposed to 17°C for four days and then to an additional recovery exposure of 14°C (below referred to as recovery exposure group). Each treatment thus consisted of four replicates, represented by 1 L Weck jars, containing 50 adult individuals at the beginning of the experiment (Figure 1). Animals were fed on the first and third day of the experiment with 10,000 cells/mL in an 8-to-2 ratio of *R. salina* and *P. cordatum* respectively. At day five, all animals were sampled from each Weck jar for mRNA extraction.

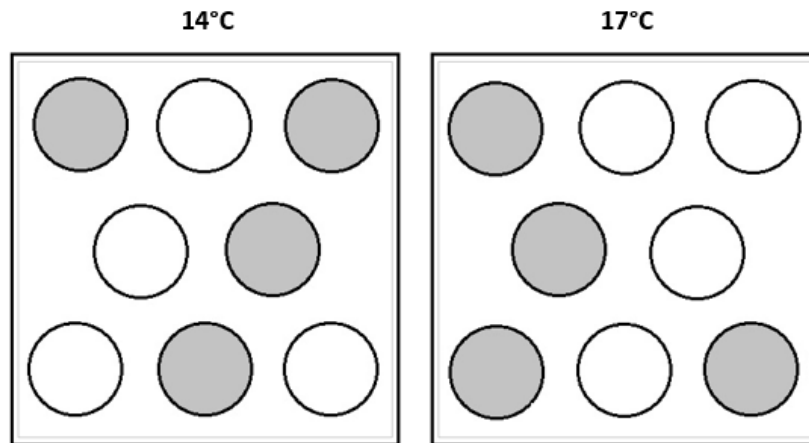


Figure 1: The set-up of the experiment: two incubators (one at 14 °C and one at 17 °C) with each eight Weck jars with 1 L of filtered sea water with an algae concentration of 10 000 algae/mL and 50 adult *T. longicornis* individuals. On day four of the experiment, four random selected jars (marked in grey) were switched randomly from one incubator to the other. Each Weck jar was aerated rigorously by an aeration pump.

2.3. RNA extraction, library preparation and sequencing

Total RNA was extracted from adult copepods using the MasterPure™ RNA Purification Kit (Epicentre®, tebu-bio NV), following the manufacturer's protocol. Remaining DNA was removed by adding RNase-Free DNase I (5 µL diluted in 195 µL DNase Buffer) (MasterPure™ RNA Purification Kit). RNA Quality was assessed with the NanoDrop™ spectrophotometer (Version 2000c, Thermo Scientific™, Thermo Fisher Scientific) (quality standards were based on the following absorbance ratios 260/230 and 260/280). The amount of RNA in the replicates was determined by means of a Qubit™ fluorometer (Version 2.0, Invitrogen™, Thermo Fisher Scientific). Exactly 10 µL of three replicates per treatment was provided to NXTGNT lab (Faculty of Pharmaceutical Sciences, UGent) for library preparation and next generation sequencing. An Illumina mRNA sequencing library was made from one µg total RNA of each sample using the TruSeq Stranded mRNA Sample Prep Kit (Illumina, San Diego, USA) with 13 PCR cycles during the enrichment PCR and according to manufacturer's protocol. The size distribution, purity (absence of free adaptors) and quantity of the resulting libraries were measured using a High Sensitivity DNA chip (Agilent Technologies, Santa Clara, CA, US). The libraries were equimolarly pooled and sequenced in three lanes of an Illumina HiSeq 3000 flow cell, generating paired-end 2x150bp reads. After sequencing, the data was demultiplexed using the sample specific nucleotide barcodes.

2.4. Sequence assembly and gene annotation

First, the quality of the raw data was assessed with FastQC Version 0.11.5 (Babraham Bioinformatics, 2010). Next, the sequences were subjected to adapter trimming and quality filtering (low-quality bases from ends of the reads) using TrimGalore Version 0.4.1 (Babraham Bioinformatics, 2012) with a threshold of an ASCII+33 Phred quality score and a length threshold of 50 bp. The remaining clean reads were then de novo assembled to form a draft transcriptome of *T. longicornis*, using the Trinity software Version 2.1.1 (Grabherr et al., 2011) with default parameters and a minimum contig length of 200 bp. Transcriptome assembly accuracy and completeness was analysed using BUSCO (Benchmarking Universal Single-Copy Orthologs; Simão et al., 2015) against a set of 843 metazoan genes and a set of 2,675 arthropod genes to evaluate the quality of the final assembly. BUSCO is a recognized benchmark approach for single copy orthologs providing an assessment of orthologs conserved among species. All transcripts were BLASTX aligned against the NCBI non-redundant (nr) arthropod protein database using BLAST2GO PRO (Conesa et al., 2005) to evaluate sequence similarity to genes in other species at the e-value cutoff of 1E-03. We also blasted (blastn) these transcripts, at the e-value cutoff of 1E-03, against three other crustacean transcriptomes, more specifically those of the calanoid copepods *Eurytemora affinis* (Almada and Tarrant, 2016), *Calanus finmarchicus* (Lenz et al., 2014) and as outgroup, the cladoceran *Daphnia magna* (Orsini et al., 2016). We also blasted the *Eurytemora* and *Calanus* transcripts to the *D. magna* transcriptome for comparison purposes.

Next, the expressed transcripts were translated into potential proteins according to ORF prediction by TransDecoder v5.0.2 (Grabherr et al., 2011). The potential coding transcripts sequences were annotated against the UniProt database (e-value of 1E-03) and InterProScan using Blast2GO PRO Annotation (Conesa et al., 2005; Jones et al., 2014). Translated contigs (peptide sequences) from the resulting functional transcriptome were also annotated with GhostKOALA (Kanehisa et al., 2015) using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The draft transcriptome has been deposited at DDBJ/EMBL/GenBank/NCBI SRA under the accession GGQN00000000. The version described in this paper is the first version, GGQN01000000.

2.5. Differential expressed Gene (DEG) analysis & Enrichment analysis of DEGs

The reference transcriptome index was constructed using Bowtie v2.3.2 (Langmead and Salzberg, 2012) using the constructed draft transcriptome as described in 2.4, and the clean paired-end reads of each sample were aligned to the reference transcriptome using TopHat v2.1.1 (Pertea et al., 2013). BAM files from TopHat and the GTF file from Cuffmerge v2.2.1 (Trapnell et al., 2010, 2013) were used as inputs for HTSeq v0.6.1 (Anders et al. 2015) to calculate gene counts with the “union” model using the Htseq-count function. These count files and the Bioconductor edgeR (v3.12.1) package (Robinson and Oshlack 2010; Robinson and Smyth 2008; Robinson et al. 2010) were used to normalize the gene counts, to estimate dispersion and to identify the differentially expressed transcripts. Normalization was performed using the “calcNormFactors” function (Robinson and Oshlack, 2010). Only transcripts with an average number counts per million reads (cpm) per sample higher than 1 were considered for DEG analysis. Three DEG analyses were performed: (1) the ‘sustained exposure’ group vs. the control group that stayed at 14 °C for five days, (2) the acute exposure group vs. the control group, and (3) the recovery exposure group vs. the ‘sustained exposure’ group. The differentially expressed transcripts were determined using a likelihood ratio test with the “glmLRT()” function and p-values were adjusted using the Benjamini-Hochberg false discovery rate (FDR) to account for multiple hypothesis testing (Benjamini and Hochberg, 1995). For the identification of significantly differentially expressed genes, only genes with FDR <0.05 were considered. These DEGS were blasted against the NCBI database (BlastX).

DEG results were then filtered based on a set of threshold values: log2FoldChange and an a p-value < 0.05. Gene ontology (GO) enrichment analysis on this filtered data set was conducted using the topGO bioconductor package in R (version 2.30.0) (Alexa and Rahnenfuhrer, 2016; Alexa et al., 2016). Enrichment analysis for GO terms against the GO annotation file from the UniProt/InterProScan database was performed exploiting the Fisher’s exact test with the weight algorithm to obtain P values (Fisher, 1922). A P-value of ≤ 0.01 was considered as a threshold for significant enrichment of the genes. Raw data files and count files were deposited in NCBI GEO under the accession number GSE115472.

3. Results

3.1 Transcriptome Assembly and Annotation

Twelve libraries of *T. longicornis* were successfully obtained from the Illumina NextSeq® 500 platform, which represented over 60 Gb of data after quality control. Clean sequencing reads

and alignment statistics are shown in Table S1. Transcriptome assembly of the Illumina reads by Trinity generated 179,569 transcripts with an average length of 540 bp (Figure 2, Table 1). Half of these transcripts (N50) were at least 694 bp long, while the longest transcript was 20,976 bp long (Table 1). Approximately 78% (660) of the BUSCO complete orthologs from the metazoan reference group were identified within the transcriptome, with 207 (24%) fragmented BUSCOs and 453 (54 %) complete BUSCOs. Among the transcripts, TransDecoder identified 44,985 putative ORF transcripts that contain candidate coding regions without a stop codon in the nucleotide sequences, resulting in a translated peptide assembly (Table 2).

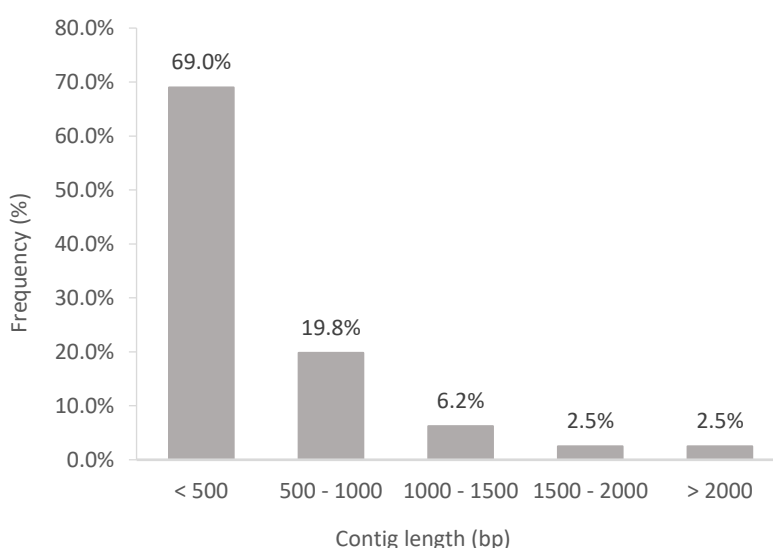


Figure 2. Frequency distribution of the contig size

Table 1. Summary statistics for the de novo assembly of the *Temora longicornis* draft transcriptome. %GC is the percent of nucleotide bases in sequences that are either G or C. N50 and N90 bp are the mean number of basepairs in all transcripts that, ordered by length, make up respectively 50% and 90% of the assembly.

Assembly statistics	
Total number of trimmed and high quality raw reads	460,460,205
Total number of assembled contigs	179,569
Total number of assembled bases (bp)	96,958,655
Minimum assembled contig length (bp)	201
Average assembled contig length (bp)	540
Median assembled contig length (bp)	334
Maximum assembled contig length (bp)	20,976

Total GC count (bp)	40,047,936
GC count for the whole assembly (%)	41.30
N50 assembled transcripts length (bp)	694
N90 assembled transcripts length (bp)	253

Table 2. Summary statistics for the annotation of the de novo assembly of the *Temora longicornis* draft transcriptome. Numbers and percentages between brackets are reported.

Annotation of transcripts encoding proteins		
Transcripts with BLAST hits against arthropod nr database		60,809
Transcripts with BLAST hits against <i>Calanus</i> transcriptome		52,004
Transcripts with BLAST hits against <i>Eurytemora</i> transcriptome		68,385
Transcripts with BLAST hits against <i>Daphnia</i> transcriptome		24,295
Putative number ORF contigs after TransDecoder		44,985
Transcripts with GO terms (UniProt & InterProScan)		20,525 (45.6%)
Transcripts with KEGG terms		14,973 (33.3%)
Number of core eukaryotic genes – BUSCO (Metazoa)		660 (78%)
	Complete genes	453 (54%)
	Complete duplicated	82 (9%)
	Fragmented genes	207 (24%)
	Missing genes	183 (22%)
Number of core eukaryotic genes – BUSCO (Arthropoda)		1676 (63%)
	Complete genes	1209 (45%)
	Complete duplicated	199 (7%)
	Fragmented genes	467 (18%)
	Missing genes	999 (37%)

BLAST searching of the assembly by sequence homology returned 60,809 peptide transcripts (33.9 %) with hits from the NCBI non-redundant (nr) arthropod protein database (Table 2; S1). Blasting the transcripts against the *Eurytemora* transcriptome revealed 68,385 different *Temora* transcripts (38% of the *T. longicornis* transcriptome), corresponding with 91,933 blast hits (covering 66.3% of the *Eurytemora* transcriptome). 29% (52,004 transcripts) the *Temora* transcriptome blast back to 75,465 sequences (37%) sequences of the complete *Calanus* transcriptome (Table 2). In comparison, 40% of the *Eurytemora* transcriptome (56,416 unique transcripts) could be blasted successfully against the *Calanus* transcriptome (75,438 blast hits

in total, also covering 37% of the entire *Calanus* transcriptome). Finally, 13.5% of the *Temora* transcriptome (24,295 sequences) displayed homology ($< e\text{-value of } 1E\text{-}03$) when aligned against the *D. magna* transcriptome (Table 2). In comparison, 14.8% (30,519) of the *Calanus* transcripts and 19% (29,129) of the *Eurytemora* transcripts could be aligned with the *D. magna* transcriptome.

Several complementary approaches were used to annotate the translated assembled sequences. In total, 65,756 (2,588 unique) GO terms and 7,875 unique IPS terms were assigned to 20,525 genes (45,6%), being annotated against UniProt (Table 2). The majority of these GO terms (Figure 3) were assigned to molecular functions (MF, 31,140 hits), followed by biological processes (BP, 18,213 hits) and cellular components (CC, 9,339 hits). A total of 33.3% (14,973) of the predicted proteins by TransDecoder were assigned to orthologous groups in the KEGG database using the GhostKoala mapping tool (Table 2, S2, S3). Most of these assigned orthologs were found to be connected to metabolic pathways (S3). Supplementary figure 4 provides a representation of the global functionality of the genes and summarizes the molecular pathways identified from the functional transcriptome. The five most represented categories of pathways were 1) Signal transduction, 2) Metabolic pathways, 3) Immune system, 4) Cell growth and death categories, and, 5) Transport and catabolism. In total, 23,423 transcripts (~52%) were not annotated by any of these tools.

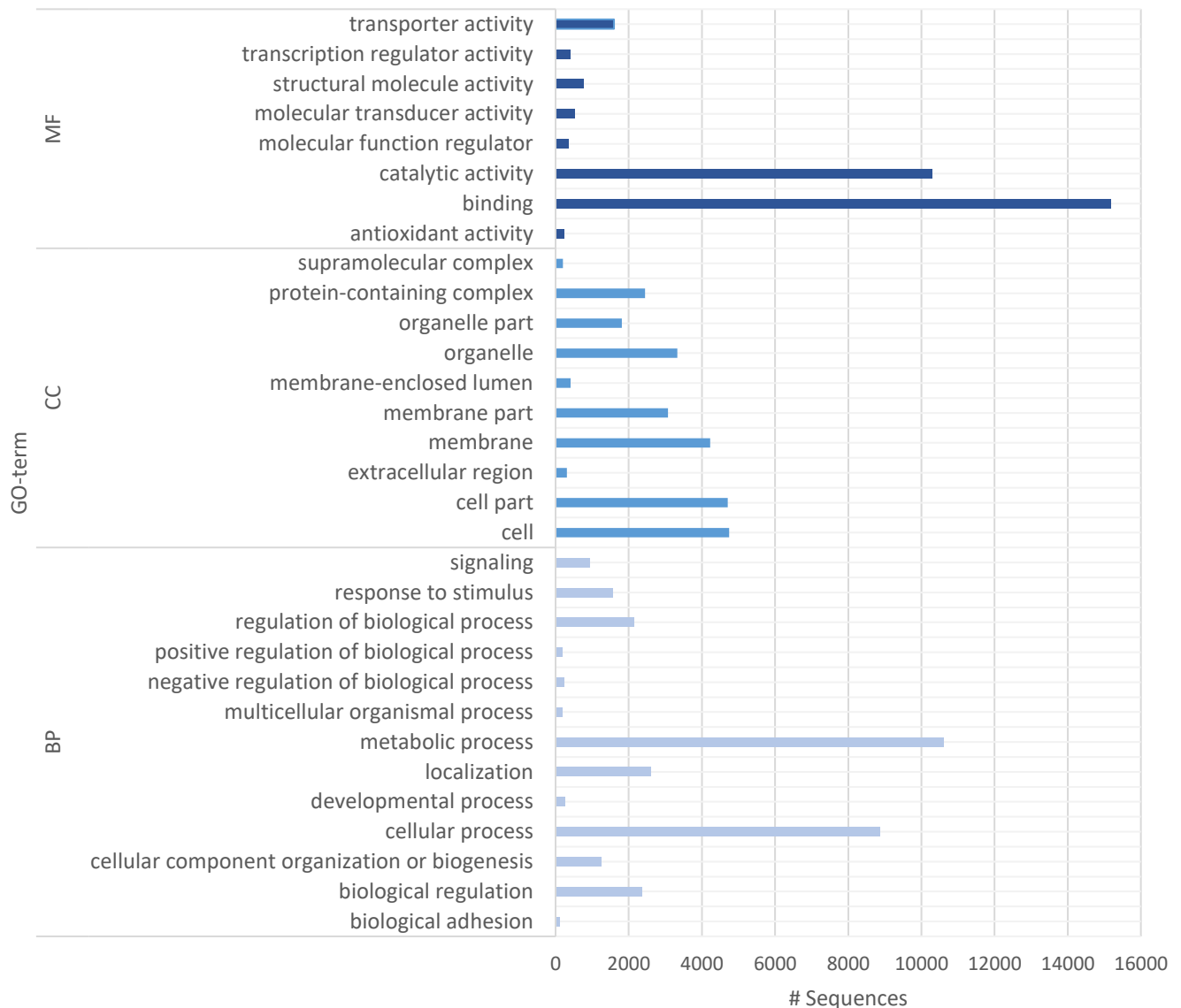


Figure 3. GO term distribution of BLAST hits from the *T. longicornis* transcriptome. Selected level 2 GO categories are shown within the divisions of Biological Process (BP), Molecular Function (MF), and Cellular Component (CC).

3.2 Gene expression under temperature

For differential expression analysis, only putative ORF transcripts with a cpm > 1 were included, resulting in 31,180 transcripts for further DEG analysis. We identified 12 differentially expressed transcripts at a FDR of 0.05 in copepods exposed to a constant temperature of 17°C (sustained heat treatment) when compared with the control group, exposed to a constant temperature of 14°C (Suppl. Table 2). The numbers of up-regulated and down-regulated transcripts in the sustained heat treatment were 5 and 7 respectively (Suppl. Table 2). Comparing the control group that was exposed at 14 °C for five days with the acute treatment

group (switched from 14 °C to 17 °C for 24 h), 15 differentially expressed transcripts were detected with 5 transcripts being down-regulated and 10 transcripts up-regulated (Suppl. Table 3). 8 transcripts were found to be significantly up- (5) or down-regulated (3) when comparing the sustained treatment group that stayed at 17 °C for five days versus the recovery group, that switched back from 17 °C to 14 °C on day 4 (Suppl. Table 4). Among all the differentially expressed genes, 20 transcripts were successfully blasted against the NCBI database (BlastX) (Suppl. Table 2, 3, 4). However, the function of most of these transcripts remains elusive (hypothetical proteins). Some of the genes were found to be affected by several treatments (Figure 4, 5): 5 transcripts were differentially expressed in both the acute and the sustained treatment, as depicted in the reaction norm plot (Figure 5, Suppl. Table 2, 3). Finally, one transcript was down-regulated in the sustained treatment, while being up-regulated in the recovery treatment (Suppl. Table 2, 4).

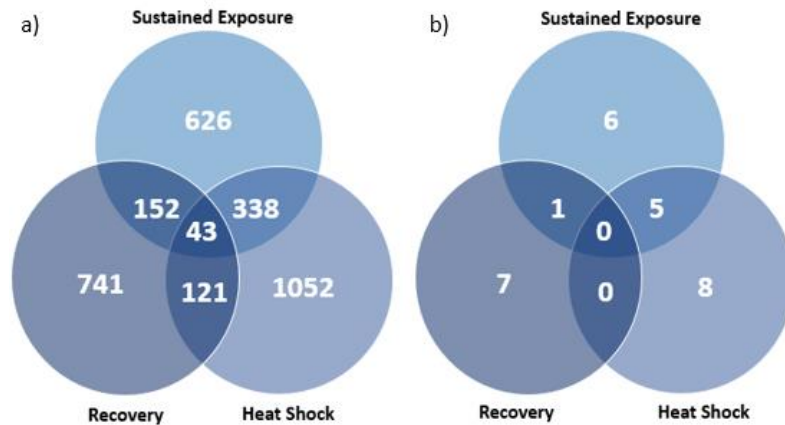


Figure 4: Gene expression comparisons of the different treatments: a Venn diagram of the number of DEGs before (a) and after (b) correcting for multiple hypothesis testing. Genes in overlapping sets show the differential expression in two or three comparison pairs.

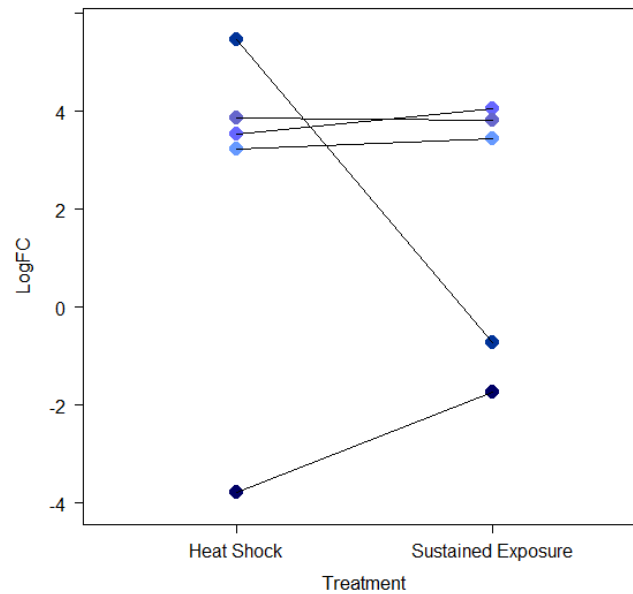


Figure 5: Thermal reaction norms of the acute heat shock treatment (Switched S17) and the sustained treatment (Continuous C17) comparing the 5 shared DEGs with the control group (Continuous C14). A steeper slope in a reaction norm is a larger log2-fold change.

Next, we conducted GO enrichment analyses across the different exposures to identify biological functions and mechanisms involved in the thermal stress response. Gene ontology analysis revealed 48 enriched GO terms in the sustained treatment (Table S5). Enrichment indicated that DEGs were mostly associated with cellular components and biological processes, with ribosome related terms being the most prominent. In molecular function, most of the annotated unique sequences were assigned to the structural constituent of ribosome and cytoskeleton, mRNA binding and GTP binding. GO term analysis revealed that DE genes in the acute treatment were strongly related to the GO terms “structural constituent of cuticle” (MF), “macromolecular complex” (CC) and “extracellular exosome” (CC) (Table S6). In total, 26 enriched GO terms were found. Finally, gene ontology analysis of the recovery treatment revealed 23 enriched GO terms (Table S7). The most enriched GO terms were related to the extracellular region for the category CC and the exoskeleton (chitin binding and structural constituent of chitin-based cuticle) for MF.

Discussion

4.1 Transcriptome assembly and comparison with other assembled copepod transcriptomes

The present study presents the first draft transcriptome of *T. longicornis*. The number of assembled transcripts by Trinity and the length of the longest transcript are comparable with other de novo copepod transcriptomes reported (Table 3). The smaller average sequence length and N50 value in our study can be partially attributed to a difference in the minimum sequence length parameter among the studies (200 bp in the present study versus 300 bp in some other studies; cf. Table 3). As the N50 metric is a commonly used metric in almost all other transcriptome studies, we chose to report this value for comparison as well. However, a N50 is not as useful in a transcriptome assembly as in a genome assembly (O’Neal and Emrich, 2003). Many metrics such as N50 and overall assembly size were simply repurposed from whole-genome assembly evaluation (O’Neal and Emrich, 2003). The N50 is used to quantify the number of large scaffolds or chromosomes as expected in a genome assembly (O’Neal and Emrich, 2003). A transcriptome however is a collection of genes, which by nature can differ significantly in length and will never form complete scaffolds or chromosomes. Another way to assess the quality of the *Temora longicornis* transcriptome is to perform a reference-based alignment using BUSCO. The BUSCO results were compared against the arthropod transcriptome reference scores provided in the BUSCO supplementary materials and several recently published crustacean transcriptomes (Figure 6) (Northcutt et al., 2016, Seong et al., 2016, Theissinger et al., 2016; Waiho et al., 2017). Our results are comparable with other recent crustacean transcriptomes such as those of *Homarus americanus* and *Cancer borealis*, but are lower compared to other copepod transcriptomes, of which BUSCO results are known (Figure 6). Overall, the annotation statistics suggest that our draft transcriptome is comparable in terms of quality and depth as other copepod transcriptomes.

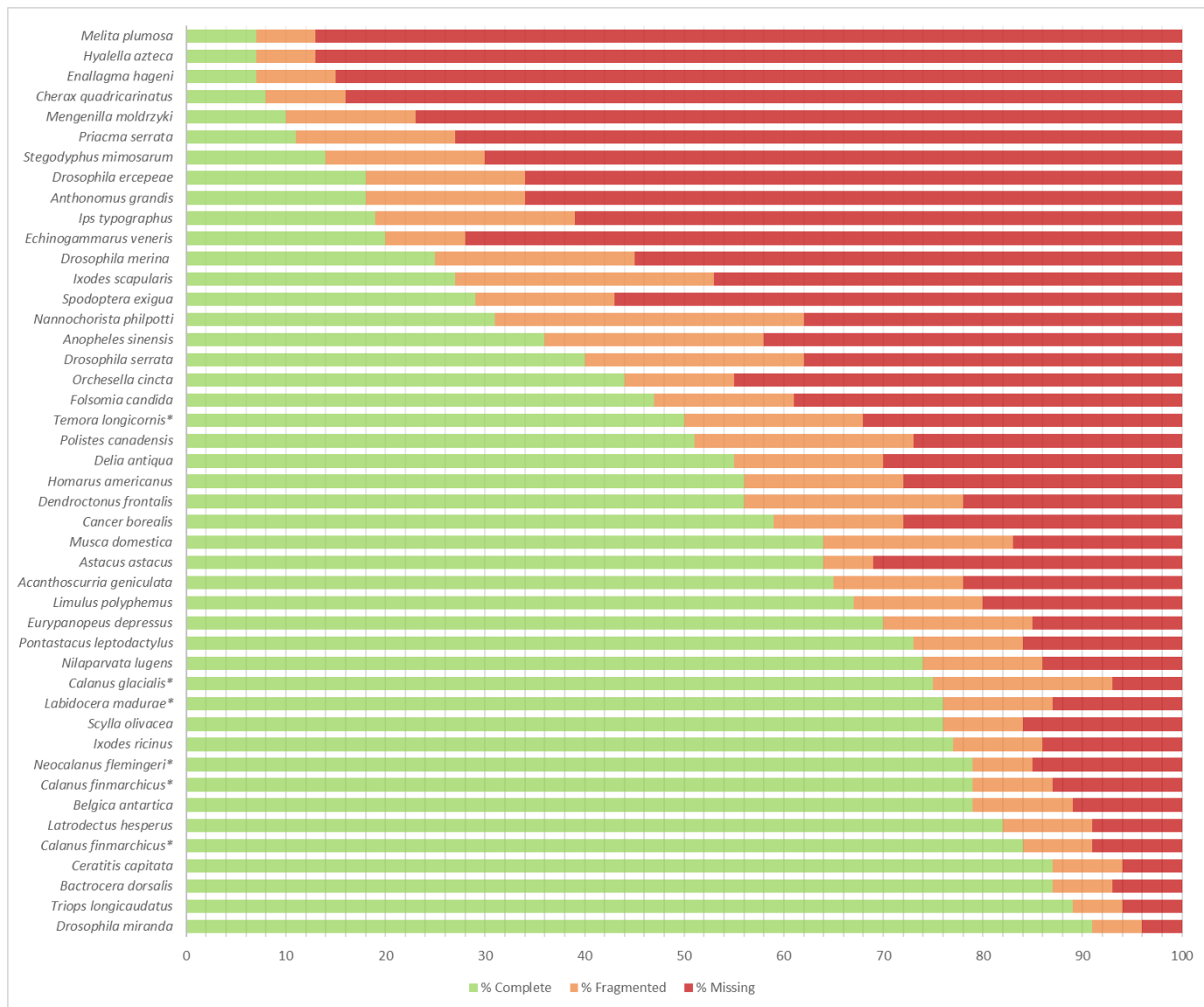


Figure 6: Horizontal stacked bar plots showing the proportions of gene sets in BUSCO quality categories for several arthropod transcriptomes. Copepod transcriptomes are denoted by a '*' (Lenz et al., 2014; Tarrant et al., 2014; Bailey et al., 2017; Roncalli et al., 2017a; Roncalli et al., 2017b). Both annotations against metazoan (M) and arthropod (A) dataset for the *Temora* transcriptome have been included. Several other recently published crustacean transcriptomes are also provided for comparison (Northcutt et al., 2016, Seong et al., 2016, Theissinger et al., 2016; Waiho et al., 2017). Quality categories are as follows: 1) Complete BUSCOs: genes that match a single gene in the BUSCO reference group (green); 2) Fragmented BUSCOs: genes only partially recovered with gene length exceeding alignment length cut-off (orange); 3) Missing BUSCOs: non-recovered genes (red).

After a successful *de novo* assembly of the RNA-seq reads, 21,562 (~48%) of the transcripts present in the *T. longicornis* draft transcriptome could be annotated with a KEGG and/or a GO term. While this may seem a low percentage, it is comparable to other non-model crustacean

species (incl. other copepod species), with a similar number of annotated genes, ranging from 20 to 55% (Table 3; Bailey et al., 2017; Kim et al., 2016; Ning et al., 2013; Roncalli et al., 2017a). De novo transcriptome studies in poorly annotated taxonomic groups such as Crustacea continue to be challenging because of these difficulties with annotation. This is due to the lack of information available on the genes of interest in closely related organisms. Only limited information exists for this taxon. To date (September 2018), the Ensembl genome database for metazoan species contains mainly insects and nematodes (Kersey et al., 2018). However, data on only three crustacean species is available: for the water fleas, *Daphnia pulex* and *Daphnia magna*, and the salmon louse, *Lepeophtheirus salmonis*, a specialised parasitic copepod species (Kersey et al., 2018). In addition, while these data are available, these genomes are also far from fully annotated. For the *Daphnia* spp. genome in particular, 35 to 45% of the coding sequences remain unannotated (Colbourne et al., 2011; Orsini et al., 2016). Furthermore, the number of GenBank records shows a threefold under-representation of crustaceans compared to insects. In total, there are 7,751,892 nucleotide sequences linked to the subphylum Crustacea, while 24,891,737 sequences are retrieved for Insecta (insects) to date (September 2018). Genomic resources available for crustaceans are hence limited, yet they are crucially needed to improve annotation rates (Stillman et al., 2008). Until now annotations are mainly based on insect genes. However, both groups include species with widely different life strategies: insects are almost exclusively terrestrial, at least in the adult stage, while crustaceans are mostly aquatic species (Alfsnes et al., 2017; Schwentner et al., 2017; Stillman et al., 2008; Tamone and Harrison, 2015). This has profound implications for the environmental drivers and life history strategies of the taxa. Comparing genome size between insects and crustaceans, Alfsnes et al. (2017) found that genome size is clearly phylogeny-dependent for insects, reflecting primarily their life history and mode of development. For crustaceans, however, they found a weaker association between genome size and phylogeny, suggesting life cycle strategies and habitat as more important determinants. Moreover, lineage specific genes are often difficult to annotate because their function is specific to both the species itself and the environment it experiences, which might additionally explain the relatively low annotation rate in the mentioned studies (Asselman et al. 2015).

4.2 Shared responses to acute and sustained thermal stress exposure

A total of 27 transcripts, were differentially expressed across heat treatments, demonstrating a pervasive genomically mild response to temperature stress, representing ~0.09% of the genes

considered for DEG analysis. The lack of up-regulation of important stress response genes could indicate that neither one of these treatments were felt as a severe stress for the copepods. While this is indeed a low number, low percentages have also been found in other studies that used non-targeted investigations of the transcriptome in response to thermal stress: in the copepods *Calanus finmarchicus* and *Tigriopus californicus* (~0.73% and ~0.88% respectively; Schoville et al., 2012; Smolina et al., 2015), the coral *Acropora hyacinthus* (~0.40%–0.70%; Barshis et al., 2013), the Antarctic fish *Trematomus bernacchii* (~0.08%–1.00%; Huth & Place, 2016) and the sea grass *Zostera marina* L. (~0.80%; Franssen et al., 2011). The response percentages of the mentioned copepod species (*C. finmarchicus* and *T. californicus*) are a bit higher than those of *T. longicornis*. This is probably due to the fact that the former species were exposed to increased temperatures just outside their natural temperature range (15°C and 35 °C respectively), whereas *Temora* was exposed to 17 °C, a temperature still within its temperature range of 5-20 °C (Halsband-Lenk et. al., 2002; Schoville et al., 2012; Smolina et al., 2015). We aimed at identifying possible target genes for acclimation to different thermal regimes (a short and more sustained exposure). General stress responses include DNA damage sensing and repair, transcription regulation, lipid metabolism and energy metabolism and damaged protein degradation (Kültz, 2005). DEG analysis identified several types of genes related to these general stress responses (Table S2, S3), but 17 out of the 27 DEGs (63 %) remain to date unannotated.

We observed seven GO terms, “structural constituent of cuticle”, “lipid metabolic process”, “response to heat”, “axonemal microtubule”, “cytosolic small ribosomal subunit”, “polysomal ribosome” and “extracellular matrix” to be enriched in genes differentially regulated in both the acute and sustained exposure treatments, indicating common transcriptional responses, caused by temperature fluctuations at different time points. Differential regulation of the ribosome and mRNA synthesis suggest an impact of temperature on protein synthesis (Table S5, S6). The differential regulation of protein synthesis in other copepod species, as well as in other marine species such as the Pacific oyster, *Crassostrea gigas*, has already been observed multiple times after exposure to thermal stress (Meistertzheim et. al., 2007; Schoville et al., 2012; Smolina et al., 2015). The “response to heat” term was also found to be enriched in other arthropod studies (e.g. Cabrera-Brandt et al., 2014; Teets et al., 2012) after exposure to pesticides or dehydration respectively, rather than to heat stress.

Enrichment analysis also revealed an overrepresentation of genes involved with the structural constituent of the cuticle in the immediate and sustained heat stress group (Table S5, S6) and

DEG analysis revealed one cuticular DEG in the latter treatment (Table S2). Cuticle proteins are structural proteins that comprise the exoskeleton of arthropods together with chitin (Liu et al. 2010). They have been found DE after thermal exposure in several previous studies as well, e.g. in *Drosophila* species (Zhao et al., 2015), the nematode *C. elegans* (Brunquell et al., 2016) and in the copepod *Tigriopus californicus* (Lima and Willett, 2017; Schoville et al., 2012), although the direction of the expression differs between and within these studies. Comparison with the control group (Table S2), shows a general down-regulation of the cuticle gene in our study. During development, the transcription of cuticle collagens is tightly regulated with moulting and growth in arthropods (Johnstone and Barry, 1996). During moulting, a new and larger exoskeleton is produced beneath the old cuticula, which is ultimately replaced so that the animal can expand in size (Eichner et al., 2015; Hartnoll, 2001). Yet, moulting in the adult phase is suppressed in copepods, a synapomorphy for Copepoda (Khodami et al., 2017). It is known though that external factors such as food and temperature affect moulting in some species (Hartnoll, 2001). We do not know the exact function of these cuticular genes/proteins in response to thermal stress, but it is possible that cuticle associated proteins are part of the thermal response in arthropods (Lima and Willett, 2017). Some studies suggest the implication of the cuticle in the regulation of the internal body temperature of arthropods (Amore et al., 2017; Carrascal et al., 2017).

Genes involved in mobilization of energy reserves (e.g. enriched GO term “lipid metabolic process”) also appear to be essential components of the general stress response in *T. longicornis*. It is known as a species with high metabolic rates but low energy reserves (Helland et al. 2003; Kreibich et. al. 2008). DEG analysis seems to support this, as it revealed a significant up-regulation for transcripts encoding for a subunit of NADH dehydrogenase (also called NADH-ubiquinone oxidoreductase), after exposure to both treatments, i.e. the acute and the more sustained period of thermal stress (Table S2, S3). A significant increase of an enzyme of the respiratory chains possibly leads to an increase in mitochondrial activity to meet an increased energy demand. This hypothesis seems to be supported by the slight up-regulation comparing the immediate and the sustained heat treatment (Figure 5). An up-regulation of this enzyme was also found in the related great spider crab *Hyas araneus* in response to ocean acidification and warming (Harms et al., 2014). In larvae of the mosquito species, *Aedes aegypti*, exposed to xenobiotics, NADH dehydrogenase was also found to be up-regulated for example, confirming this DEG is being part of a global stress response (David et al., 2010).

One annotated transcript related to damage control, apolipoprotein D-like isoform, was also DE in both heat treatments (exposure to acute and a more sustained period of thermal stress) (Table S2, S3). Apolipoprotein D (ApoD) and isoforms are ancient members of the lipocalin family with a high degree of sequence conservation from plants to insects to mammals (Ganfornina et al., 2008; Muffat and Walker, 2010). ApoD has emerged as an evolutionarily conserved anti-stress protein, by maintaining low levels of lipid peroxidation, that is induced by both internal and external stressors and inflammation (Ganfornina et al., 2008; Muffat and Walker, 2010). ApoD was found to be up-regulated after induction to several types of stressors in studies on human cells, mammals, fruit flies and plants (Charron et al., 2008; Dassati et al., 2014; Muffat and Walker, 2010). Interestingly, in our study we showed that the gene coding for the ApoD isoform was down-regulated after exposure to thermal stress, especially after exposure to acute stress, with a small increase after exposure to more sustained stress (Table S2, S3). In the study of Xu et al. (2016) two down-regulated apolipoprotein D genes were identified upon cold stress in the white-backed planthopper, *Sogatella furcifera*.

4.3 Signatures of acute thermal stress

“Macromolecular complex”, “extracellular exosome”, “short-chain fatty acid catabolic process” and “glyceraldehyde-3-phosphate dehydrogenase (NAD⁺) (phosphorylating) activity” are the most significant GO terms enriched in genes differentially regulated by acute heat stress (Table S5, S6). These GO terms were not significantly enriched in genes up-regulated after a more sustained exposure to heat stress, suggesting a temporal specificity of these responses to heat stress. The Go term “chitin binding” though, can be linked to cuticle related GO terms, a response shared between the exposure scenarios (discussed in 4.2). Eight transcripts were DE specifically after exposure to short term heat stress as well, including the down-regulated gene, ganglioside GM2 activator-like. GM2 activators act within lysosomes, where they present GM2 gangliosides to a coenzyme to be broken down (d’Azzo and Sano, 2006). Several crucial cellular responses may be influenced or controlled by gangliosides and ultimately result in either cell growth and division, differentiation, or cell death (d’Azzo and Sano, 2006; Tessitore et al., 2004). Under stress conditions that dramatically increase their intracellular concentration, gangliosides can initiate the induction of an apoptotic program (d’Azzo and Sano, 2006; Tessitore et al., 2004). Thus, a down-regulation of enzymes that break down gangliosides, may possibly lead to an increased abundance of gangliosides, suggesting perceived stress at the cellular level.

Several enriched GO terms (e.g. “short-chain fatty acid catabolic process”, “carbohydrate phosphorylation” and “glyceraldehyde-3-phosphate dehydrogenase”) are associated with energy metabolism. We also found an increase in the expression of the SLC37A2-like gene (also known as glucose-6-phosphate transporter/exchanger, G6PT) (Table S3). Even though there is still little information on its exact protein function, it is described to transport glucose-6-phosphate, a key component of glycolysis, the first fundamental metabolic process of aerobic respiration (Chou and Mansfield, 2014; Olcese and Bartolini, 2013). A possible explanation of the up-regulation of this gene is the additional requirement for energy of the organism as a general stress response. Respiration rates of several marine calanoid copepod species, including *T. longicornis* and several *Centropages* species, were indeed found to increase with temperature following an exponential function (Castellani and Altunbas, 2013; Cruz et al., 2013; Gaudy, 1973, 1975). In another study, Vidal (1980) found a linear increase of the metabolic rate with a logarithmic increase in temperature in the marine calanoid copepod *Calanus pacificus*. These findings, together with our results, indicate indeed an increased energy demand after exposure to increased water temperatures.

4.4 Signatures of sustained thermal stress

Overall, the transcriptome of *T. longicornis* demonstrated a larger response to a sustained heat exposure, rather than an immediate (acute) exposure, with two times as many enriched GO terms in the sustained treatment. Gene ontology analysis indicated that the differentially expressed genes were enriched with a broad spectrum of biological processes and pathways (48 GO terms in total) among which protein synthesis appears to be the most affected metabolic pathway. After exposing *Calanus finmarchicus* copepods to realistic, sustained thermal stress (+ 5°C), Smolina et al., (2015) also found the enrichment of the GO terms “structural constituent of the ribosome” and “cytosolic small ribosomal subunit” (cf. our results, Table S5), suggesting an effort to increase translation capacity or to protect the ribosomal function through the replacement/addition of ribosomal proteins (Kültz, 2005). Several GO terms, such as “protein binding”, “mRNA binding”, “nucleosome” and “structural constituent of cytoskeleton”, were found to be significantly enriched in genes differentially regulated after a more sustained exposure only (Table S5). Nucleosome remodelling has already been shown to be an important part of the stress response, e.g. by allowing access to transcription sites of stress responsive genes (de Nadal et al., 2011; Guertin et al., 2010; Jovic et al., 2017). Depletion of nucleosome remodelling leads to a higher thermal sensitivity in *C. elegans* for example (Satyal et al., 1998). Hence, packaging of DNA into nucleosomes could be an additional protective mechanism

during the stress response (Jovic et al., 2017). In addition to nucleosome remodelling, genes encoding components of the cytoskeleton were also found to be enriched after a sustained exposure to heat stress (Table S5). Cytoskeletal organization is known to go through a profound transformation under thermal stress (Richter et al., 2010). The GO term “structural constituent of cytoskeleton” was also found as enriched in other heat stress exposure studies in for example trypanosomatids and the Hong Kong Oyster, *Crassostrea hongkongensis* (Kelly et al., 2012; Li et al., 2017). Again, several enriched GO terms, related to metabolism and energy (e.g. “ATP synthesis coupled proton transport”, Table S5), suggest that *T. longicornis* individuals experienced a higher metabolic rate when submitted to a sustained heat temperature treatment.

We also identified one GO term (“heat shock protein binding”; interacting selectively and non-covalently with a heat shock protein) associated with heat-shock proteins (Hsps) in the sustained exposure treatment (Table S5). The same GO term was also found to be significantly enriched with a lower pH in *Calanus glacialis* (Bailey et al., 2017). A general cellular reaction to a stressor is the expression of molecular chaperones (De Maio, 1999; Moseley, 1997). As molecular chaperones, heat-shock proteins are involved in the folding and unfolding of other proteins with proteins that are in their non-native conformation (stress denatured) in such a way that they prevent these proteins from interacting inappropriately with each other (De Maio, 1999; Moseley, 1997). Heat shock proteins are typically increasing in abundance when cells are exposed to elevated temperatures or other stressors (De Maio, 1999; Moseley, 1997). Through BLAST searching the nr protein database, a total of 171 Hsps and 31 Hsfs (heat-shock factor) were identified in the functional transcriptome, none of which were significantly differentially expressed (p-value <0.05; Table S5). In addition to their chaperone activity, hsps also target damaged proteins to the proteasome to prevent accumulation of dysfunctional proteins and to recycle amino acids (Goldberg, 2003). We detected the enrichment of a GO term related to proteolysis (Table S5) in the DE genes. The lack of an abundant number of differentially expressed chaperone genes after both an acute and a sustained treatment in our study could indicate that (1) only a limited number of proteins is affected by the increase of temperature to 17°C, making an up-regulated expression of Hsps unnecessary, or (2) that up-regulation of the Hsps happened already a few hours (less than 24 hours) after the treatments and was already back to normal regulation when RNA was extracted.

4.5 Signatures of recovery after thermal stress

To assess whether *T. longicornis* individuals could recover from exposure to water temperatures of 17° C, we compared gene expression between two groups exposed to 17°C, with one of them having had a day of recovery exposure of 14°C on day 4 (the recovery treatment). We found some enriched GO terms related to energy and protein metabolism (e.g. “alpha,alpha-trehalase activity”, Table S7). Interestingly, the GO terms “chitin binding”, “structural constituent of the chitin-based cuticle”, “structural constituent of the extracellular matrix” and “chaperone-mediated protein folding” were also enriched, confirming their role in the stress response of *T. longicornis*. DE analysis revealed only 8 DEGs. One gene, trypsin-1 like (Table S4), was shown to be down-regulated after recovery. Trypsins, having a proteolytic function, have been established as key enzymes for growth through its role in the protein digestion processes (Rungruangsak Torrisen and Male 2000). Moreover, trypsin is recognized as the most abundant protease present in crustaceans (Jones et al. 1997). Trypsins have already been reported being differentially expressed, in both directions, after exposure to stressors such as cyanobacteria, cadmium and salinity stress in the crustacean *Daphnia magna* (Asselman et al., 2012; Heckman et al., 2008; Poynton et al., 2007). The removal of a stressor might be an explanation for the down-regulation in our experiment. Simultaneously, DEG analysis revealed the up-regulation of a part of the protease inhibitor inter-alpha-trypsin inhibitors (ITI) heavy chain H3-like, suggesting indeed an inhibition of trypsin activity. On the other hand, there is strong evidence that all members of the ITI family play an important role in extracellular matrix stability by covalent linkage to hyaluronan, linking the ITI H3-like gene with the GO term “structural constituent of the extracellular matrix” (Hamm et al., 2008). Recovery from stress is dependent on several factors, such as exposure duration, timing of stress application in terms of development, population densities and other ecological conditions, as has been shown in several studies on *Daphnia* (Barata et al., 2002; Hanazato and Dodson, 1995; Reynaldi and Liess, 2005). While this study presents a first effort to reveal the possibility of stress recovery in *Temora*, more detailed research is necessary to understand the recovery potential of this species from stress.

4.6 Linking molecular responses to potential higher-level effects

Our results from the DEG and GO enrichment analyses suggest *Temora longicornis* may display a general cellular stress response when exposed to a water temperature increase of 3°C. More research is needed to confirm the downstream biological and developmental effects of the differential expression of these genes. The effect of temperature-induced gene expression

changes on the fitness of *Temora* in the BPNS remains thus to be addressed. Our findings suggest a weak response to exposure to a temperature of 17 °C, which might be interpreted as an indication that *T. longicornis* is naturally tolerant to this temperature, which is consistent with the occurrence of this species above the 20 °C isotherm in summer (Halsband-Lenk et. al., 2002). Moreover, using ad libitum feeding conditions, Holste et al. (2009) measured egg production of *T. longicornis* in the Baltic Sea at 12 different temperatures between 2.5 and 24°C and observed the highest mean egg production at 16.9°C (12 eggs female⁻¹ day⁻¹), suggesting a tolerance to an increased temperature of the sea water. The findings of Holste et al. (2009) contrast with Halsband-Lenk et. al. (2002), who did not find large differences in reproduction rates over a temperature range of 2-20 °C. Deschutter et al. (2017) applied multimodel inference on generalized additive models to quantify the relative contribution of temperature to the dynamics of calanoid copepod species in the Belgian part of the North Sea. For *Temora*, they found that temperature is consistently important for the population density of this species: an optimal temperature between 14°C and 16.5°C, together with negative effects below temperatures of ± 11 °C was found for this species. Together, these studies indicate a tolerance to water temperatures of 17°C.

Although a GO enrichment analysis offers initial insights into processes affected by an elevated temperature, a strong bias still exists towards conserved and well-characterized molecular functions and cellular components in model organisms (Harms et al., 2014). Species-specific transcripts are not assessed when they are not annotated and may thus form a bottleneck for the identification of key stressor pathways. This bias particularly applies to non-model organisms, such as *Temora*, with a lack of GO annotation for about half of the transcripts. Additionally, many genes are grouped into more than one GO term and are thus more difficult to interpret. Considering these limitations, a GO enrichment analysis can only provide a general overview of possibly affected processes, making more detailed analyses desirable. Also, to predict whether *T. longicornis* will be able to cope with the temperature rise - due to global warming - more elaborate experiments with more extreme temperature stress treatments should be carried out (instead of mild temperature exposure treatments only). In addition, it could also be interesting to assess gender-specific and life-stage specific differences in the molecular stress response of *Temora*. Boulangé-Lecomte et. al. (2014) found for example evidence for sexual dimorphism in the heat shock protein expression (Grp78 and Hsp90A) in the closely related estuarine copepod *Eurytemora affinis*, as females showed significantly higher expression for the tested Hsp genes than male individuals, suggesting females might be more stress resistant.

Finally, transcriptomic responses may not necessarily predict changes at the protein level. Limitations of monitoring exclusively gene expression, such as posttranslational modifications and the destabilising mRNA under stress, makes mRNA an inaccurate proxy for the abundance of active products of translation (Androulakis et al., 2007). We acknowledge that mRNA transcript abundance is indeed not equivalent to protein abundance and that ultimately it is the stressed proteome that protects and repairs the cell during stressor exposure. Yet, other studies have shown though that there is often a high correlation between gene expression and protein abundance and that gene expression changes in experimental treatments are often relevant predictors of organismal physiology (Feder and Walser, 2005; Maier et al., 2009). Hence, future exposure studies should further try to link gene expression data with phenotypic data.

Conclusions

The first transcriptome analysis on the calanoid copepod *Temora longicornis* was carried out here. However, the annotation of the gene transcripts proved to be more difficult due to the lack of available information or closely related genomes. The implementation of whole transcriptome technology to study the response to a thermal stressor in *T. longicornis* has resulted in identifying several general stress responses independent of exposure time, such as modified protein synthesis, energy mobilisation, cuticle and chaperone proteins. Overall, *T. longicornis* was more impacted by a sustained exposure, rather than an immediate (acute) exposure, to thermal stress with two times as many enriched GO terms in the sustained treatment. Furthermore, we also highlighted several transcripts and enriched GO terms with a temporal specificity, illustrating the importance of exposure time in gene transcription. Finally, we highlighted candidate genes of a possible recovery from heat exposure, identifying similar terms as those enriched in the heat treatments, i.e. related to for example energy metabolism, cuticle genes and extracellular matrix. In conclusion, *Temora* exhibits a general, mild cellular stress response when exposed to water temperature increase of 3°C. However, more research is needed to confirm the downstream biological, developmental and fitness effects of these differentially expressed transcripts.

Acknowledgements

The authors are grateful to the Flanders Marine Institute (VLIZ) for the assistance, especially André Cattijse and the crew of the research vessel Simon Stevin. The computational resources (Stevin Supercomputer Infrastructure) and services used in this work were provided by the VSC

(Flemish Supercomputer Center), funded by Ghent University, FWO and the Flemish Government – department EWI. The authors thank Josefien Van Landuyt, Pieter Masschelein and Olivier Berteloot for the technical assistance. Funding was received from UGent research fund (BOF15/24J/106). Jana Asselman is an FWO (Fonds voor Wetenschappelijk Onderzoek – Vlaanderen) postdoctoral fellow.

References

Alexa, A. and Rahnenfuhrer, J. (2016). topGO: Enrichment Analysis for Gene Ontology. R package version 2.30.0.

Alexa, A., Rahnenfuhrer, J., and Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating go graph structure. *Bioinformatics* (Oxford, England), 22: 1600–1607.

Alfsnes, K., Leinaas, H. P., & Hessen, D. O. (2017). Genome size in arthropods; different roles of phylogeny, habitat and life history in insects and crustaceans. *Ecology and Evolution*, 7(15): 5939–5947.

Almada, A. A., Tarrant, V. (2016). *Vibrio* elicits targeted transcriptional responses from copepod hosts. *FEMS Microbiol. Ecol.*, 92 (6).

Amore, V., Hernández, M. I. M., Carrascal, L. M., & Lobo, J. M. (2017). Exoskeleton may influence the internal body temperatures of Neotropical dung beetles (Col. Scarabaeinae). *PeerJ*, 5, e3349.

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166-169.

Androulakis, I. P., Yang, E., & Almon, R. R. (2007). Analysis of Time-Series Gene Expression Data: Methods, Challenges, and Opportunities. *Annual Review of Biomedical Engineering*, 9, 205–228.

Aral, M. M. and Guan, J. (2016). Global Sea Surface Temperature and Sea Level Rise Estimation with Optimal Historical Time Lag Data. *Water* 8: 519.

Asselman, J., De Coninck, D., Glaholt, S., Colbourne, J. K., Janssen, C., Shaw, J. R., & De Schampelaere, K. A. C. (2012). Identification of pathways, gene networks, and paralogous gene families in *Daphnia pulex* responding to exposure to the toxic cyanobacterium *Microcystis aeruginosa*. *Environmental Science & Technology*, 46(15): 8448–8457.

Asselman, J., Pfrender, M. E., Lopez, J. A., De Coninck, D., Janssen, C., Shaw, J. R., & De Schampelaere, K. (2015). Conserved transcriptional responses to cyanobacterial stressors are mediated by alternate regulation, of paralogous genes in *Daphnia*. *Molecular Ecology*, 24(8): 1844–1855.

729 Atkinson, D. (1994). Temperature and organism size — a biological law for ectotherms. *Adv.*
730 *Ecol. Res.* 25: 1–58.

731 d’Azzo, A. T. A. & Sano, R. (2006). Gangliosides as apoptotic signals in ER stress response.
732 *Cell Death Differ.*, 13: 404–414.

733 Bailey, A., De Wit, P., Thor, P., Browman, H. I., Bjelland, R., Shema, S., Fields, D. M., Runge,
734 J. A., Thompson, C., Hop, H. (2017). Regulation of gene expression is associated with tolerance
735 of Artic copepod *Calanus glacialis* to CO₂-acidified water. *Ecol. Evol.* In press.

736 Barata C., Baird, D.J., Soares, A.M. (2002). Demographic responses of a tropical cladoceran to
737 cadmium: Effects of food supply and density. *Ecol. Appl.*, 12: 552–564.

738 Barreto, F.S., Moy, G.W., Burton RS. (2011). Interpopulation patterns of divergence and
739 selection across the transcriptome of the copepod *Tigriopus californicus*. *Molecular Ecology*,
740 20: 560–572.

741 Beaugrand, G. and Kirby, R. R. (2010). Climate, plankton and cod. *Global Change*
742 *Biology*, 16: 1268–1280.

743 Beaugrand, G., Edwards, M., Brander, K., Luczak, C., Ibanez, F. (2008). Causes and projections
744 of abrupt climate-driven ecosystem shifts in the North Atlantic. *Ecology Letters* 11: 1157–1168.

745 Beaugrand, G., Luczak, C., Edwards, M. (2009). Rapid biogeographical plankton shifts in the
746 North Atlantic Ocean. *Global Change Biology* 15: 1790–1803.

747 Beaugrand, G., Reid, P. C., Ibanez, F., Lindley, J. A., Edwards, M. (2002). Reorganization of
748 North Atlantic marine copepod biodiversity and climate. *Science* 296: 1692–1694.

749 Babraham Bioinformatics: *FastQC A Quality Control tool for High Throughput Sequence*
750 *Data*. (2010). – URL: <http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/>. Date accessed:
751 2017-06-08.

752 Babraham Bioinformatics: *Babraham Bioinformatics – Trim Galore!* (2012). – URL:
753 http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/. Date accessed: 2017-06-
754 12.

755 Barshis, D. J., Ladner, J. T., Oliver, T. A., Seneca, F. O., Traylor-Knowles, N., & Palumbi, S.
756 R. (2013). Genomic basis for coral resilience to climate change. *Proceedings of the National*
757 *Academy of Sciences of the United States of America*, 110: 1387–1392.

758 Benjamini, Y.; Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and
759 Powerful Approach to Multiple Testing. *J Roy Stat Soc B Met.* 57, (1): 289-300.

760 Boulangé-Lecomte, C., Forget-Leray, J., Xuereb, B. (2014). Sexual dimorphism in Grp78 and
761 Hsp90A heat shock protein expression in the estuarine copepod *Eurytemora affinis*. *Cell stress*
762 *& chaperones* 19 (4): 591–7.

763 Bron, J. E., Frisch, D., Goetze, E., Johnson, S. C., Lee, C. E. and Wyngaard, G. A. (2011).
764 Observing copepods through a genomic lens. *Frontiers in zoology* 8 (1): 22.

765 Brunquell J., Morris S., Lu Y., Cheng F., Westerheide S.D. (2016). The genome-wide role of
766 HSF-1 in the regulation of gene expression in *Caenorhabditis elegans*. *BMC Genomics*. 17:
767 559.

768 Brylinski, J. M. (2009). The pelagic copepods in the Strait of Dover (eastern English Channel).
769 A commented inventory 120 years after Eugène Canu. *Cahiers de Biologie Marine* 50: 251-
770 260.

771 Cabrera-Brandt, M., Silva, A. X., Le Trionnaire, G., Tagu, D., and Figueroa, C. C. (2014).
772 Transcriptomic responses of the aphid *Myzus persicae nicotianae* Blackman (Hemiptera:
773 Aphididae) to insecticides: Analyses in the single Chilean clone of the tobacco aphid. *Chilean*
774 *journal of agricultural research*, 74(2), 191-199.

775 Carrascal, L.M., Jiménez-Ruiz, Y., Lobo, J.M. (2017). Beetle exoskeleton may facilitate body
776 heat acting differentially across the electromagnetic spectrum. *Physiological and Biochemical*
777 *Zoology* 90:338–347.

778 Casini, M., Cardinale, M., Arrhenius, F. (2004). Feeding preferences of herring (*Clupea*
779 *harengus*) and sprat (*Sprattus sprattus*) in the southern Baltic Sea. *ICES Journal of Marine*
780 *Science* 61 (8): 1267–1277.

781 Castellani, C., & Altunbaş, Y. (2013). Seasonal change in acclimatised respiration rate of
782 *Temora longicornis* (Müller). *Marine Ecology Progress Series*, 500. 10.3354/meps10661.

783 Charron, J. B., Ouellet, F., Houde, M., Sarhan, F. (2008). The plant Apolipoprotein D ortholog
784 protects *Arabidopsis* against oxidative stress. *BMC Plant Biol.*, 8: 86.

785 Chiba, S., Batten, S., Martin, C. S., Ivory, S., Miloslavich, P. and Weatherdon, L. V. (2018).
786 Zooplankton monitoring to contribute towards addressing global biodiversity conservation
787 challenges, *Journal of Plankton Research*, fby030.

788 Chou, J. Y., Mansfield, B.C. (2014). The SLC37 family of sugar-phosphate/phosphate exchangers. *Curr. Top. Membr.*, 73: pp. 357-382.

789 Colbourne, J. K.; Pfrender, M. E.; Gilbert, D.; Thomas, W. K.; Tucker, A.; Oakley, T. H.;
790 Tokishita, S.; Aerts, A.; Arnold, G. J.; Basu, M. K.; Bauer, D. J.; Caceres, C. E.; Carmel, L.;
791 Casola, C.; Choi, J. H.; Detter, J. C.; Dong, Q.; Dusheyko, S.; Eads, B. D.; Frohlich, T.; Geiler-
792 Samerotte, K. A.; Gerlach, D.; Hatcher, P.; Jogdeo, S.; Krijgsveld, J.; Kriventseva, E. V.; Kultz,
793 D.; Laforsch, C.; Lindquist, E.; Lopez, J.; Manak, J. R.; Muller, J.; Pangilinan, J.; Patwardhan,
794 R. P.; Pitluck, S.; Pritham, E. J.; Rechtsteiner, A.; Rho, M.; Rogozin, I. B.; Sakarya, O.;
795 Salamov, A.; Schaack, S.; Shapiro, H.; Shiga, Y.; Skalitzy, C.; Smith, Z.; Souvorov, A.; Sung,
796 W.; Tang, Z.; Tsuchiya, D.; Tu, H.; Vos, H.; Wang, M.; Wolf, Y. I.; Yamagata, H.; Yamada,
797 T.; Ye, Y.; Shaw, J. R.; Andrews, J.; Crease, T. J.; Tang, H.; Lucas, S. M.; Robertson, H. M.;
798 Bork, P.; Koonin, E. V.; Zdobnov, E. M.; Grigoriev, I. V.; Lynch, M.; Boore, J. L. The
799 ecoresponsive genome of *Daphnia pulex*. (2011). *Science*, 331 (6017), 555– 61.

800 Conesa, A., Götz S., García-Gómez, J. M., Terol, J., Talón, M. and Robles, M. (2005).
801 Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics
802 research. *Bioinformatics*, 21(18): 3674–3676.

803 Crain, C. M., Halpern, B. S., Beck, M.W., Kappel, C. V. (2009). Understanding and managing
804 human threats to the coastal marine environment. *Ann. NY Acad. Sci.* 1162: 39–62.

805 Daan, R. (1989). Factors controlling the summer development of the copepod populations in
806 the southern bight of the North Sea. *Neth. J. Sea Res.* 23: 305–322.

807 Dam, H.G. and Peterson, W.T. (1993). Seasonal contrast in the diel vertical distribution, feeding
808 and grazing impact of the copepod *Temora longicornis* in Long Island Sound. *J. Mar. Res.* 51:
809 561–594.

810 Dassati, S., Waldner, A., Schweigreiter, R. (2014). Apolipoprotein D takes center stage in the
811 stress response of the aging and degenerative brain. *Neurobiol Aging*, 35: 1632-42.

812 David, J.-P., Coissac, E., Melodelima, C., Poupardin, R., Riaz, M. A., Chandor-Proust, A.
813 (2010). Transcriptome response to pollutants and insecticides in the dengue vector *Aedes*
814 *aegypti* using next-generation sequencing technology. *BMC Genomics*, 11: 216.

815 De Maio, A. (1999). Heat shock proteins: facts, thoughts, and dreams. *Shock* 11: 1–12.

816 de Nadal, E., Ammerer, G., Posas, F. (2011). Controlling gene expression in response to stress.
817 *Nat Rev Genet. Nature Publishing Group*, 12: 833–45.

818 Deschutter, Y., Everaert, G., De Schamphelaere, K. A. C. and De Troch, M. (2017). Relative
819 contribution of multiple stressors on copepod density and diversity dynamics in the Belgian
820 part of the North Sea. *Marine Pollution Bulletin*, 125(1-2), 350–359.

821 Devreker, D., Souissi, S., Seuront, L. (2005). Effects of chlorophyll concentration and
822 temperature variation on the reproduction and survival of *Temora longicornis* (Copepoda,
823 Calanoida) in the Eastern English Channel. *Journal of Experimental Marine Biology and*
824 *Ecology* 318 (2): 145–162.

825 Dippner, J. W., Kornilovs, G., Sidrevics, L. (2000). Long-term variability of mesozooplankton
826 in the Central Baltic Sea. *J. Mar. Sys.* 25: 23-31.

827 Dzierzbicka-Glowacka, L., Lemieszek, A. and Żmijewska, M. I. (2011). Development and
828 growth of *Temora longicornis*: numerical simulations using laboratory culture data.
829 *Oceanologia* 53(1): 137–161.

830 Eichner, C., Hamre, L. A., Nilsen, F. (2015). Instar growth and molt increments in
831 *Lepeophtheirus salmonis* (Copepoda: Caligidae) chalimus larvae. *Parasitol Int.* 2015;64: 86–
832 96. pmid:25451218

833 Feder, M.E., Walser, J. C. (2005). The biological limitations of transcriptomics in elucidating
834 stress and stress responses. *J Evol Biol.*, 18: 901–910.

835 Ferdous Z and Muktadir A. (2009). A Review: Potentiality of Zooplankton as Bioindicator.
836 *American Journal of Applied Sciences*, 6(10): 1815–9.

837 Fisher, R. A. (1922). On the interpretation of χ^2 from contingency tables, and the calculation
838 of P. *J. R. Stat. Soc.*, 85: 87–94.

839 Franssen, S. U., Gu, J., Bergmann, N., Winters, G., Klostermeier, U. C., Rosenstiel, P., Reusch,
840 T. B. H. (2011). Transcriptomic resilience to global warming in the seagrass *Zostera marina*, a
841 marine foundation species. *Proceedings of the National Academy of Sciences*, 108: 19276–
842 19281

843 Gallardo-Escárate, C., Valenzuela-Muñoz, V., & Nuñez-Acuña, G. (2014). RNA-Seq Analysis
844 Using De Novo Transcriptome Assembly as a Reference for the Salmon Louse *Caligus*
845 *rogercresseyi*. *PLoS ONE*, 9(4), e92239.

846 Ganfornina, M. D., Do Carmo, S., Lora, J. M., Torres-Schumann, S., Vogel, M., Allhorn, M.,
847 Sanchez, D. (2008). Apolipoprotein D is involved in the mechanisms regulating protection from
848 oxidative stress. *Aging Cell*, 7(4): 506–515.

849 Gaudy, R. (1973). Les variations saisonnières de la respiration chez quatre espèces de
850 copépodes pélagiques du Golfe de Marseille. *Neth. J. Sea Res.*, 7: 267-279.

851 Gaudy, R. (1975). Etude de la respiration chez des copépodes pélagiques méditerranéens
852 (bassin occidental et mer Ionienne) et de ses variations en fonction de la bathymétrie des espèces
853 et de leur origine géographique. *Marine Biology*, 29: 109-118.

854 Goldberg, A.L. (2003). Protein degradation and protection against misfolded or damaged
855 proteins. *Nature* 426 (6968):895–899.

856 Gonçalves, R. J., van Someren Gréve, H., Couespel, D. and Kiørboe, T. (2014). Mechanisms
857 of prey size selection in a suspension-feeding copepod, *Temora longicornis*. *Mar. Ecol. Prog.*
858 *Ser.* 517: 61–74.

859 Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis,
860 X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A.,
861 Rhind, N., Palma, F.d., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman and N., Regev,
862 A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference
863 genome. *Nature biotechnology* 29 (7): 644–52.

864 Guertin, M. J., Petesch, S. J., Zobeck, K. L., Min, I. M., Lis, J.T. (2010). *Drosophila* heat shock
865 system as a general model to investigate transcriptional regulation. *Cold Spring Harb Symp*
866 *Quant Biol.*, 75: 1–9.

867 Halpern, B. S., Walbridge, S., Selkoe, K. A., Kappel, C.V., Micheli, F., D'Agrosa, C., Bruno,
868 J.F., Casey, K.S., Ebert, C., Fox, H.E. (2008). A global map of human impact on marine
869 ecosystems. *Science* 319: 948-952.

870 Halsband-Lenk, C., Hirche, H. -J. and Carlotti, F. (2002). Temperature impact on reproduction
871 and development of congener copepod populations. *J. Exp. Mar. Biol. Ecol.*, 271: 121–153.

872 Hamm, A., Veeck, J., Bektas, N., Wild, P. J., Hartmann, A., Heindrichs, U. Dahl, E. (2008).
873 Frequent expression loss of Inter-alpha-trypsin inhibitor heavy chain (ITI_H) genes in multiple
874 human solid tumors: A systematic expression analysis. *BMC Cancer*, 8: 25.

875 Hanazato, T., Dodson, S.I. (1995). Synergistic effects of low oxygen concentration, predator
876 kairomone, and a pesticide on the cladoceran *Daphnia pulex*. *Limnol. Oceanogr.*, 40:700–709.

877 Harms, L., Frickenhaus, S., Schiffer, M., Mark, F. C., Storch, D., Held, C., Lucassen, M. (2014).
878 Gene expression profiling in gills of the great spider crab *Hyas araneus* in response to ocean
879 acidification and warming. *BMC Genomics*, 15(1): 789.

880 Hartnoll, R.G. (2001). Growth in crustacea — twenty years on *Hydrobiologia*, 449, 111-122.

881 Hay, S. J., Evans, G. T. and Gamble J. C. (1988). Birth, growth and death rates for enclosed
882 populations of calanoid copepods. *J. Plankton Res.* 10 (3): 431–454.

883 Heckmann, L.-H., Sibly, R. M., Connon, R., Hooper, H. L., Hutchinson, T. H., Maund, S. J.,
884 and Callaghan, A. (2008). Systems biology meets stress ecology: linking molecular and
885 organismal stress responses in *Daphnia magna*. *Genome Biology*, 9(2): R40.

886 Helland, S., Terjesen, B. F., Berg, L. (2003). Free amino acid and protein content in the
887 planktonic copepod *Temora longicornis* compared to *Artemia franciscana*. *Aquaculture* 215,
888 1-4, p213–228.

889 Heyen, H., Fock, H. and Greve, W. (1998). Detecting relationships between the interannual
890 variability in ecological timeseries and climate using a multivariate statistical approach – a case
891 study on Helgoland Roads zooplankton. *Clim. Res.* 10: 179-191.

892 Holste, L., St. John, M. A., Peck, M. A. (2009). The effects of temperature and salinity on
893 reproductive success of *Temora longicornis* in the Baltic Sea: A copepod coping with a tough
894 situation. *Marine Biology* 156 (4): 527–540.

895 Huth, T. J., & Place, S. P. (2016). Transcriptome wide analyses reveal a sustained cellular stress
896 response in the gill tissue of *Trematomus bernacchii* after acclimation to multiple stressors.
897 *BMC Genomics*, 17: 1–18.

898 IPCC (Intergovernmental Panel on Climate Change). *Climate Change 2007: The Physical*
899 *Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the*
900 *Intergovernmental Panel on Climate Change*. Solomon, S., Qin, D., Manning, M., Chen, Z.,
901 Marquis, M., Averyt, K. B., Tignor, M., Miller, H. L., Eds. Cambridge University Press,
902 Cambridge, UK; New York, NY, USA, 2007.

903 IPCC (Intergovernmental Panel on Climate Change). *Climate Change 2014: Synthesis Report.*
904 *Contribution of Working Groups I, II and III to the Fifth Assessment Report of the*
905 *Intergovernmental Panel on Climate Change*. Pachauri, R. K., Meyer, L. A., Core Writing
906 Team, Eds. Geneva, Switzerland: IPCC. 2014. 151 p.

907 Cruz, J., Garrido, S., Pimentel, M. S., Rosa, R., Santos, A. M. P., Ré, P. (2013). Reproduction
908 and respiration of a climate change indicator species: effect of temperature and variable food in
909 the copepod *Centropages chierchiae*. *Journal of Plankton Research*, 35 (5): 1046–1058.

910 Johnstone, I. L. and Barry, J. D. (1996). Temporal reiteration of a precise gene expression
911 pattern during nematode development. *EMBO J. European Molecular Biology Organization*,
912 15: 3633–9.

913 Jones, D. A., Kumlu, M., Le Vay, L., Fletcher, D. J. (1997). The digestive physiology of
 914 herbivorous, omnivorous, and carnivorous crustacean larvae: a review. *Aquaculture*, 155: 285–
 915 295.

916 Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W. Z., Mcanulla, C., Hunter, S. (2014).
 917 InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, 30, 1236–1240

918 Jovic, K., Sterken, M. G., Grilli, J., Bevers, R. P. J., Rodriguez, M., Riksen, J. A. G., Snoek, L.
 919 B. (2017). Temporal dynamics of gene expression in heat-stressed *Caenorhabditis elegans*.
 920 *PLoS ONE*, 12(12), e0189445.

921 Kane, J. and Prezioso, J. (2008). Distribution and multi-annual abundance trends of the copepod
 922 *Temora longicornis* in the US Northeast Shelf Ecosystem. *J. Plankton Res.* 30 (5): 619-632.

923 Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG
 924 tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* 428,
 925 726-731.

926 Kelly, S., Kramer, S., Schwede, A., Maini, P. K., Gull, K., & Carrington, M. (2012). Genome
 927 organization is a major component of gene expression control in response to stress and during
 928 the cell division cycle in trypanosomes. *Open Biology*, 2(4), 120033.

929 Kersey, P. J., Allen, J. E., Allot, A., Barba, M., Boddu, S., Bolt, B. J., Carvalho-Silva, D.,
 930 Christensen, M., Davis, P., Grabmueller, C., Kumar, N., Liu, Z., Maurel, T., Moore, B.,
 931 McDowall, M. D., Maheswari, U., Naamati, G., Newman, V., Ong, C. K., Bolser, D. M., De
 932 Silva, N., Howe, K. L., Langridge, N., Maslen, G., Staines, D. M., Yates, A. (2018). Ensembl
 933 Genomes 2018: an integrated omics infrastructure for non-vertebrate species. *Nucleic Acids*
 934 *Research* 46(D1) D802–D808.

935 Khodami, S., McArthur, J. V., Blanco-Bercial, L., & Martinez Arbizu, P. (2017). Molecular
 936 Phylogeny and Revision of Copepod Orders (Crustacea: Copepoda). *Scientific Reports*, 7, 9164.
 937 <http://doi.org/10.1038/s41598-017-06656-4>

938 Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, SL. (2013) TopHat2:
 939 accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions.
 940 *Genome Biol.* 14: R36.

941 Kim, H.-S., Lee, B.-Y., Won, E.-J., Han, J., Hwang, D.-S., Park, H.G., Lee, J.-S. (2015).
 942 Identification of xenobiotic biodegradation and metabolism-related genes in the copepod
 943 *Tigriopus japonicus* whole transcriptome analysis. *Marine Genomics*, Part 3 24, 207-208.

944 Kim, H. -S., Lee, B. -Y., Han, J., Lee, H. Y, Min, G.-S., Kim, S., Lee, J.-S. (2016). De novo
 945 assembly and annotation of the Antarctic copepod (*Tigriopus kingsejongensis*) transcriptome.
 946 *Marine Genomics*, 28, 37-39.

947 Klein Breteler, W. C. M. (1980). Continuous breeding of marine pelagic copepods in the
 948 presence of heterotrophic dinoflagellates. *Mar. Ecol. Prog. Ser.*, 2, pp. 229-233.

949 Klein Breteler, W. C. M. and Gonzalez, S. R. (1986). Culture and development of *Temora*
950 *longicornis* (Copepoda, Calanoida) cultured at different temperature and food conditions, *Mar.*
951 *Ecol. Prog. Ser.*, 119: 99–110.

952 Kreibich, T., Saborowski, R., Hagen, W., Niehoff, B. (2008). Short-term variation of nutritive
953 and metabolic parameters in *Temora longicornis* females (Crustacea, Copepoda) as a response
954 to diet shift and starvation. *Helgoland Marine Research* 62, 3: 241–249.

955 Klimaatinfo: *Zeeland weer & klimaat - alles over het klimaat van Zeeland*. 2016. – URL:
956 <http://www.klimaatinfo.nl/nederland/zeeland.htm>. – Date accessed: 2016-04-24.

957 Krause, M., Dippner, J. W., Beil, J. (1995). A review of hydrographic controls on the
958 distribution of zooplankton biomass and species in the North Sea with particular reference to a
959 survey conducted in January–March 1987. *Progress in Oceanography* 35: 81-152.

960 Kültz, D. (2005). Molecular and evolutionary basis of the cellular stress response. *Annu. Rev.*
961 *Physiol.*, 67: 225–257.

962 Langmead, B. and Salzberg, S. (2012). Fast gapped-read alignment with Bowtie 2. *Nature*
963 *Methods* 9: 357-359.

964 Lee, B. -Y., Kim, H. -S., Choi, B. -S., Hwang, D.-S., Choi, A. Y., Han, J., Won, E. -J., Choi, I.-
965 Y., Lee, S. -H., Om, A.-S., Park, H. G. and Lee, J.-S. (2015). RNA-seq based whole
966 transcriptome analysis of the cyclopoid copepod *Paracyclopina nana* focusing on xenobiotics
967 metabolism. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 15:
968 12–19.

969 Lenz, P.H., Roncalli, V., Hassett, R.P., Wu, L.S., Cieslak, M.C., Hartline, D.K. and Christie,
970 A.E. (2014). De novo assembly of a transcriptome for *Calanus finmarchicus* (crustacea,
971 copepoda) - The dominant zooplankter of the North Atlantic Ocean. *PLoS ONE* 9 (2).

972 Li, J., Zhang, Y., Mao, F., Tong, Y., Liu, Y., Zhang, Y and Yu, Z. (2017). Characterization and
973 identification of differentially expressed genes involved in thermal adaptation of the Hong
974 Kong oyster *Crassostrea hongkongensis* by digital gene expression profiling. *Front Mar Sci.*
975 4: 112.

976 Lima, T. G., & Willett, C. S. (2017). Locally adapted populations of a copepod can evolve
977 different gene expression patterns under the same environmental pressures. *Ecology and*
978 *Evolution*, 7(12), 4312–4325.

979 Liu, Q., Yuan, Y., Lin, J., Zhong, Y. (2010). Advance of researches on insect cuticular proteins
980 and the regulation mechanism of their gene expression (in Chinese with English abstract).
981 *Chinese Journal of Applied Entomology*, 2: 247–255.

982 Maier, T., Güell, M., Serrano, L. (2009). Correlation of mRNA and protein in complex
983 biological samples. *FEBS Lett.*, 583: 3966–3973.

984 Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing
985 reads. *EMBnet.journal* 17 (1): 10.

986 McCarthy, F. M., Wang, N., Magee, G. B., Nanduri, B., Lawrence, M. L., Camon, E. B., Barrell,
 987 D. G., Hill, D. P., Dolan, M. E., Williams, W. P., Luthe, D. S., Bridges, S. M., Burgess, S. C.
 988 (2006). AgBase: A functional genomics resource for agriculture. *BMC Genomics*. 7:229.

989 McNeil, B., and R. Matear (2006). Projected climate change impact on oceanic acidification,
 990 *Carbon Balance and Manage.*, 1(2).

991 Meistertzheim, A. –L., Tanguy, A., Moraga, D., Thébault, M. –T. (2007). Identification of
 992 differentially expressed genes of the Pacific oyster *Crassostrea gigas* exposed to prolonged
 993 thermal stress. *FEBS Journal* 274 (24): 6392–6402.

994 Möllmann, C., Kornilovs, G., Fetter, M., Koster, F. W. (2005). Climate, zooplankton, and
 995 pelagic fish growth in the central Baltic Sea. *ICES J. Mar. Sci.* 62: 1270–1280.

996 Moseley, P.L. (1997). Heat shock proteins and heat adaptation of the whole organism. *J. Appl.*
 997 *Physiol.* 83: 1413–1417.

998 Muffat, J., & Walker, D. W. (2010). Apolipoprotein D: An overview of its role in aging and
 999 age-related diseases. *Cell Cycle* (Georgetown, Tex.), 9(2): 269–273.

1000 Nagelkerken, I. & Connell, S. D. (2015). Global alteration of ocean ecosystem functioning due
 1001 to increasing human CO₂ emissions. *Proc. Natl. Acad. Sci. USA* 112, 13272–13277.

1002 Ning, J., Wang, M., Li, C., Sun, S. (2013). Transcriptome Sequencing and De Novo Analysis
 1003 of the Copepod *Calanus sinicus* using 454 GS FLX. *PLoS ONE* 8 (5).

1004 Northcutt, A. J., Lett, K. M., Garcia, V. B., Diester, C. M., Lane, B. J., Marder, E. (2016). Deep
 1005 sequencing of transcriptomes from the nervous systems of two decapod crustaceans to
 1006 characterize genes important for neural circuit function and modulation. *BMC Genomics*
 1007 17:868.

1008 O’Neil, S. T., & Emrich, S. J. (2013). Assessing De Novo transcriptome assembly metrics for
 1009 consistency and utility. *BMC Genomics*, 14, 465. <http://doi.org/10.1186/1471-2164-14-465>

1010 Olcese, C. and Bartoloni, L. (2013), The sugar phosphate/phosphate exchanger family SLC37.
 1011 *WIREs Membr Transp Signal*, 2: 255–264.

1012 Orr, J. C., Fabry, V. J., Aumont, O., Bopp, L., Doney, S.C., Feely, R.A., Gnanadesikan, A.,
 1013 Gruber, N., Ishida, A., Joos, F., Key, R.M., Lindsay, K., Maier-Reimer, E., Matear, R.,
 1014 Monfray, P. (2005). Anthropogenic ocean acidification over the twenty-first century and its
 1015 impact on calcifying organisms, *Nature*, 437: 681–686.

1016 Orsini, L., Gilbert, D., Podicheti, R., Jansen, M., Brown, J. B., Solari, O. S., Spanier, K. I.,
 1017 Colbourne, J. K., Rush, D., Decaestecker, E., Asselman, J., De Schamphelaere, K. A. C., Ebert,
 1018 D., Haag, C. R., Kvist, J., Petrusek, A., Pfrender, M. E., De Meester, L. and Frilander, M. J.
 1019 (2016). *Daphnia magna* transcriptome by RNA-Seq across 12 environmental stressors. *Sci data*
 1020 3:160030. doi: 10.1038/sdata.2016.30.

1021 OSPAR: Quality Status report 2010 / OSPAR commission. London, UK, (2010). – Research
 1022 report. – 176 p.

1023 Pierce, D. W., Gleckler, P. J., Barnett, T. P., Santer, B. D. and Durack, P. J. (2012). The
 1024 fingerprint of human-induced changes in the ocean's salinity and temperature fields. *Geophys.*
 1025 *Res. Lett.* 39, L21704.

1026 Poynton, H. C., Varshavsky, J. R., Chang, B., Cavigliolo, G., Chan, S., Holman, P. S.,
 1027 Loguinov, A. V. (2007). *Daphnia magna* ecotoxicogenomics provides mechanistic insights into
 1028 metal toxicity. *Environ. Sci. Technol.* 41: 1044–1050.

1029 Reynaldi, S., Liess, M. (2005). Influence of duration of exposure to the pyrethroid fenvalerate
 1030 on sublethal responses and recovery of *Daphnia magna* Straus. *Environ. Toxicol. Chem.* 24:
 1031 1160–1164.

1032 Richardson, A.J. (2008). In hot water: Zooplankton and climate change. *ICES Journal of*
 1033 *Marine Science* 65 (3): 279–295.

1034 Richter, K., Haslbeck, M., and Buchner, J. (2010). The heat shock response: life on the verge
 1035 of death. *Mol Cell* 40, 253–266.

1036 Rivkin, R.B., and Legendre, L. (2001). Biogenic carbon cycling in the upper ocean: effects of
 1037 microbial respiration. *Science* 291 (5512): 2398–2400.

1038 Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential
 1039 expression analysis of RNA-seq data. *Genome Biol.* 11: R25.

1040 Robinson, M. D., and Smyth, G. K. (2008). Small-sample estimation of negative binomial
 1041 dispersion, with applications to SAGE data. *Biostatistics* 9:3 21–332.

1042 Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: A Bioconductor package
 1043 for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140.

1044 Roncalli, V., Cieslak, S. C., Sommer, R. A., Hopcroft, R., Lenz, P. H. (2017a). De novo
 1045 transcriptome assembly of the calanoid copepod *Neocalanus flemingeri*: A new resource for
 1046 emergence from diapause. *Marine Genomics*. In press.

1047 Roncalli, V., Christie, A. E., Sommer, S. A., Cieslak, M. C., Hartline, D. K., Lenz, P. H. (2017).
 1048 A deep transcriptomic resource for the copepod crustacean *Labidocera madurae*: a potential
 1049 indicator species for assessing near shore ecosystem health. *PLoS One*. (In press.)

1050 Rungruangsak Torrisen K., Male R. (2000). Trypsin isozymes: development, digestion and
 1051 structure. In: Haard N.F., Simpson B.K., editors. *Seafood Enzymes, Utilization and Influence*
 1052 *on Postharvest Seafood Quality*. New York: Marcel Dekker, Inc.; pp. 215–269.

1053 Satyal, S.H., Chen, D., Fox, S.G., Kramer, J.M., Morimoto, R.I. (1998). Negative regulation of
 1054 the heat shock transcriptional response by HSBP1. *Genes Dev. Cold Spring Harbor Laboratory*
 1055 *Press*, 12: 1962–1974.

1056 Schoville, S.D., Barreto, F.S., Moy, G.W., Wolff, A., Burton, R.S. (2012). Investigating the
 1057 molecular basis of local adaptation to thermal stress: population differences in gene expression
 1058 across the transcriptome of the copepod *Tigriopus californicus*. *BMC Evol. Biol.*, 12 (1): 170.

1059 Schwentner, M., Combosch, D. J., Pakes Nelson, J., & Giribet, G. (2017). A phylogenomic
 1060 solution to the origin of insects by resolving crustacean-hexapod relationships. *Current Biology*,
 1061 27, 1818–1824.

1062 Seong, J., Kang, S. W., Patnaik, B. B., Park, S. Y., Hwang, H. J., Chung, J. M., Song, D. K.,
 1063 Noh, M. Y., Park, S. -H., Jeon, G. J., (2016). Transcriptome analysis of the tadpole shrimp
 1064 (*Triops longicaudatus*) by Illumina paired-end sequencing: Assembly, annotation, and marker
 1065 discovery. *Genes*, 7, 114.

1066 Sharqawy, M. H., Lienhard J. H., and Zubair S. M. (2010). The thermophysical properties of
 1067 seawater: A review of existing correlations and data, *Desalin. Water Treat.*, 16: 354–380.

1068 Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M. (2015).
 1069 BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs.
 1070 *Bioinformatics* 31(19): 3210–3212.

1071 Stillman, J. H., J. K. Colbourne, C. E. Lee, N. H. Patel, M. R. Phillips, D. W. Towle, B. D.
 1072 Eads, G. W. Gelembuik, R. P. Henry, E. A. Johnson, M. E. Pfrender & N. B. Terwilliger.
 1073 (2008). Recent advances in crustacean genomics. *Integr. Comp. Biol.* 48:852–868.

1074 Tamone, S.L. and Harrison, J.F. (2015). Linking Insects with Crustacea: Physiology of the
 1075 Pancrustacea: An Introduction to the Symposium. *Integrative and Comparative Biology*, 55 (5):
 1076 765–770.

1077 Tarrant, A. M., Baumgartner, M. F., Hansen, B. H., Altin, D., Nordtug, T., Olsen, A. J. (2014).
 1078 Transcriptional profiling of reproductive development, lipid storage and molting throughout the
 1079 last juvenile stage of the marine copepod *Calanus finmarchicus*. *Front. Zool.* 11:91.

1080 Teets, N. M., Peyton, J. T., Colinet, H., Renault, D., Kelley, J. L., Kawarasaki, Y., Lee, R. E.
 1081 and Denlinger, D. L. (2012). Gene expression changes governing extreme dehydration
 1082 tolerance in an Antarctic insect. *Proc. Natl. Acad. Sci. USA* 109, 20744–20749.

1083 Tessitore, A., del P.M.M., Sano, R., Ma, Y., Mann, L., Ingrassia, A., Laywell, E. D., Steindler,
 1084 D. A., Hendershot, L. M., and d'Azzo, A. (2004). GM1-ganglioside-mediated activation of the
 1085 unfolded protein response causes neuronal death in a neurodegenerative gangliosidosis. *Mol.*
 1086 *Cell.*, 15: 753–766

1087 Theissinger, K., Falckenhayn, C., Blande, D., Toljamo A, Gutekunst J, Makkonen J, (2016).
 1088 De Novo assembly and annotation of the freshwater crayfish *Astacus astacus* transcriptome.
 1089 *Mar Genomics*. 2016;28:7–10.

1090 Trapnell, C., Hendrickson, D. G., Sauvageau, M., Goff, L., Rinn, J. L., Pachter, L. (2013).
 1091 Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.*
 1092 31(1): 46–53.

1093 Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M.J., Salzberg,
 1094 S.L., Pachter, L. (2010). Transcript assembly and quantification by RNA-seq reveals
 1095 unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28:
 1096 511–515.

1097 van der Kooij, J., Engelhard, G.H., Righton, D.A. (2016). Climate change and squid range
1098 expansion in the North Sea. *J. Biogeogr.* 43, 2285–2298.

1099 Van Ginderdeuren, K., Van Hoey, G., Vincx, M. and Hostens, K. (2014b). The
1100 mesozooplankton community of the Belgian shelf (North Sea). *J. Sea Res.* 85: 48-58.

1101 Van Ginderdeuren, K., Vandendriessche, S., Prossler, Y., Matola, H., Vincx, M., Hostens, K.
1102 (2014a) Selective feeding by pelagic fish in the Belgian part of the North Sea. *ICES Journal of*
1103 *Marine Science* 71 (4): 808–820.

1104 Vidal, J. (1980). Physioecology of zooplankton. III. Effects of phytoplankton concentration,
1105 temperature, and body size on the metabolic rate of *Calanus pacificus*. *Marine Biology*, 56 (3):
1106 195–202.

1107 Villarino, E., Chust, G., Licandro, P., Butenschön, M., Ibaibarriaga, L., Larrañaga, A. (2015).
1108 Modelling the future biogeography of North Atlantic zooplankton communities in response to
1109 climate change. *Mar. Ecol. Prog. Ser.* 531: 121–142.

1110 Waiho K., Fazhan, H., Shahreza, M.S., Moh, J. H., Noorbaiduri, S., Wong, L. L., (2017).
1111 Transcriptome Analysis and Differential Gene Expression on the Testis of Orange Mud Crab,
1112 *Scylla olivacea*, during Sexual Maturation. *Plos One*, 12(1):e0171095.

1113 Williams, R., Conway, D. V. P., Hunt, H. G. (1994). The role of copepods in the planktonic
1114 ecosystems of mixed and stratified waters of the European shelf seas. *Hydrobiologia* 292/293:
1115 521-530.

1116 Xu, D., Zhong, T., Feng, W., Zhou, G. (2016). Tolerance and responsive gene expression of
1117 *Sogatella furcifera* under extreme temperature stresses are altered by its vectored plant virus.
1118 *Sci Rep.*, 6: e31521.

1119 Zacharias, M. A., Gregr, E. J. (2005). Sensitivity and vulnerability in marine environments: an
1120 approach to identifying vulnerable marine areas. *Conserv. Biol.* 19: 86-97.

1121 Zhao L., Zhao, L., Wit J., Svetec, N., . (2015). Parallel gene expression differences between
1122 low and high latitude populations of *Drosophila melanogaster* and *D. simulans* . *PLOS*
1123 *Genetics*, 11(5): e1005184.

Table 3. Comparison of de novo assembled transcriptomes for several copepod species. Summary statistics are reported, unless not provided by the study ('nr': not reported). Where possible, summary statistics were calculated (marked with an '*'). Respective data sources: 1) Current study; annotation results of the functional draft transcriptome are included in the table; 2) Almada and Tarrant, 2016; 3) Lenz et al., 2014; 4) Tarrant et al., 2014; 5) Bailey et al., 2017; 6) Ning et al., 2013; 7) Roncalli et al., 2017a; 8) Roncalli et al., 2017b; 9) Schoville et al., 2012, San Diego population 10) Kim et al., 2015; 11) Kim et al., 2016; 12) Lee et al., 2015; 13) Gallardo-Escárate et al., 2014. Percentages are given between brackets.

	Calanoida								Harpacticoida			Cyclopoida	Siphonostomatoida
	<i>Temora longicornis</i> ¹	<i>Eurytemora affinis</i> ²	<i>Calanus finmarchicus</i> ³	<i>Calanus finmarchicus</i> ⁴	<i>Calanus glacialis</i> ⁵	<i>Calanus sinicus</i> ⁶	<i>Neocalanus flemingeri</i> ⁷	<i>Labidocera madurae</i> ⁸	<i>Tigriopus californicus</i> ⁹	<i>Tigriopus japonicus</i> ¹⁰	<i>Tigriopus king-sejongensis</i> ¹¹	<i>Paracyclopina nana</i> ¹²	<i>Caligus rogercresseyi</i> ¹³
Sequencing platform	Illumina HiSeq	Illumina HiSeq	Illumina HiSeq	Illumina HiSeq	Illumina HiSeq	454 GS FLX	Illumina NextSeq	Illumina NextSeq	Illumina HiSeq	Illumina NextSeq	Illumina HiSeq	Illumina HiSeq	MiSeq Illumina
Assembly software	Trinity	Trinity	Trinity	Trinity	CLC	Newbler	Trinity & CAP3	Trinity	CLC	Trinity	Trinity	Trinity	CLC
Raw Reads (#)	460,460,205	216,737,662	401,836,653	93,000,000	636,738,663	1,470,799	149,805,914	528,000,341	9,292,573	108,072,846	139,923,976	205,692,994	154,840,000
Assembled Transcripts (#)	179,569	138,581	206,041	241,140	nr	56,809	245,135	211,002	66,943	140,130	81,653	125,631	83,444
Transcripts with ORF (#)	44,985 (26%)	82,891 (59%)	96,090 (47%)	124,618 (51%)	59,353	/	108,092 (44%)	72,391 (34%)	/	54,761 (39%)	38,250 (49%)	67,179 (53%)	/
Minimum Length (bp)	201	201	301	201	195*	201	301	301	49	201	224	201	300
Average Length (bp)	540	857	997	667	650*	nr	1011*	872	498	1,695	954	2,260	819
Maximum Length (bp)	20,976	23,627	23,068	25,048	13,363	nr	24,981	23,836	9,830	30,174	8427	30,223	nr
N50	694	2,087	1,418	988	1,019	873	1,452	1,184	nr	3,565	1,283	4,178	nr
Transcripts with BLAST hits (#)	60,809	nr	81,554	nr	nr	9,497	62,126	62,980	17,405	39,507	22,977	12,474	23,841
Transcripts with GO terms (#)	20,525 (45.6%)	nr	78,467 (81%)	nr	19,790 (33%)	25,676 (45%)	59,544 (55%)	60,097 (83%)	nr	27,706 (50%)	16,815 (44%)	nr	15,314 (19%)
Transcripts with KEGG terms (#)	14,973 (33%)	nr	67,673 (70%)	nr	nr	8,962 (18%)	52,555 (49%)	57,912 (80%)	nr	nr	nr	nr	16,213 (20%)